Juan Pedro Milaszewicz

On Brown's method with convexity hypotheses

Applications of Mathematics, Vol. 49 (2004), No. 2, 165–184

Persistent URL: http://dml.cz/dmlcz/134565

ON BROWN'S METHOD WITH CONVEXITY HYPOTHESES

JUAN PEDRO MILASZEWICZ, Buenos Aires

(Received April 29, 2002, in revised version September 6, 2003)

Abstract. Given two initial points generating monotone convergent Brown iterations in the context of the monotone Newton theorem (MNT), it is proved that if one of them is an upper bound of the other, then the same holds for each pair of respective terms in the Brown sequences they generate. This comparison result is carried over to the corresponding Brown-Fourier iterations. An illustration is discussed.

Keywords: nonlinear systems, convex functions, Brown's method, monotone convergence, Fourier iterates

MSC 2000: 65H10

1. INTRODUCTION

For a continuously differentiable function $F\colon D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$, consider the associated system

$$F(x) = 0. \tag{1.1}$$

In order to find an approximate solution of (1.1) when $n > 1$, Brown proposed a combination of the one dimensional Newton method with a Gauss-Seidel-like extension to the nonlinear case of the Gaussian elimination process. Quadratic convergence for Brown's analytic method was established (see [2]), but the original algorithm had a cost of $O(n^4)$ algebraic operations per iteration, which made it unattractive when compared with Newton's method. Subsequently, this figure was reduced to $O(n^3)$ with a better implementation (see [1] and [11]). With this new algorithm and in the context of the MNT (see [9]), Frommer proved a monotone Brown theorem (see [3]), as well as that Brown's analytic method converges componentwise at least as fast as Newton's analytic method (see [4]); these results include the corresponding ones for the respective Fourier iterations. Comparison theorems hold for Newton's method (see [6]) in the MNT framework. Thus it is interesting and can be useful to obtain analogous results for Brown's method in the same context, and this is the aim of the present paper. A typical comparison result proved here in that framework is that, if $z^0 \leqslant y^0$ are such that $0 \leqslant F(y^0)$, $0 \leqslant F(z^0)$, and if $F(y^*) = 0$, then $y^* \leqslant z^k \leqslant y^k$ for all Brown iterates with starting values z^0 and y^0, respectively, where the inequalities are understood componentwise.

It is convenient to recall that the MNT context arises naturally when mildly nonlinear problems and certain nonlinear integral equations are discretized. For instance, when discretizing $\Delta u = \mathrm{e}^u$ by means of second central difference quotients (see [9]), and when discretizing Chandrasekhar's integral equation (see the last section). Also, if $\Delta u = \mathrm{e}^u$ is transformed, via Green's function, into an equivalent integral equation, then its discretization by means of a trapezoidal rule generates the same context again.

Although the comparison statements for Brown's method are formally similar to those for Newton's method (see [6]), they require proofs that are much more involved. In order to simplify cumbersome calculations, the main idea here consists in proving that the Brown function assigning to each initial point its first Brown iterate, is a monotone function with respect to the coordinatewise standard order on a convenient domain. This idea has also been employed for a third order method in [7] and, though the general approach in both papers is similar, the proofs regarding the two methods are essentially unrelated.

The comparison results for Brown's method are also extended to the corresponding Fourier iterations which coupled with the Brown iterations provide bracketings, i.e. lower and upper bounds of the solution of (1.1).

The outline of the paper is as follows. The next section describes the framework and some significant conclusions are drawn regarding the geometry of the zero manifolds of f_i, $1 \leqslant i \leqslant n$; the section contains the only common features with [7], while correcting a minor flaw in a proof of that paper. The third section contains the main results for Brown's analytic method, while in the fourth their extensions to the Brown-Fourier iterations are proved as well. In the fifth section, a numerical example illustrates the results.

2. Definitions and basic results

It is assumed in the sequel that $x^0 < y^0$, i.e. $x^0_i < y^0_i$, $1 \leqslant i \leqslant n$, and that

$$\langle x^0, y^0 \rangle := \{x\colon\ x^0 \leqslant x \leqslant y^0\} \subset D \quad \text{and} \quad F(x^0) \leqslant 0 \leqslant F(y^0).$$

Also for y in D, $F'(y)$ is a nonsingular M-matrix (see [12]), i.e.

$$(F'(y))_{i,j} \leqslant 0 \quad \text{for} \quad i \neq j \quad \text{and} \quad F'(y)^{-1} \geqslant 0.$$

It is also assumed that the Jacobian matrix F' is isotone, that is

$$F'(x) \leqslant F'(y) \quad \text{if} \quad x \leqslant y.$$

Recall that if F' is isotone, then F is order convex, namely

$$F(\lambda x + (1-\lambda)y) \leqslant \lambda F(x) + (1-\lambda)F(y), \quad \text{if} \quad x \leqslant y, \quad \text{and} \quad \lambda \in (0,1).$$

The MNT ensures that there exists a unique y^* in $\langle x^0, y^0 \rangle$, for which (1.1) holds (see Chapter 13 in [9]). Moreover, the Newton iterates with starting point y^0 converge decreasingly, while their Fourier counterparts with starting point x^0 do so increasingly, and the bracketing they both determine converges quadratically to 0. As a consequence, if y in $\langle x^0, y^0 \rangle$ satisfies $F(y) \geqslant 0$ ($F(y) \leqslant 0$), then $y \geqslant y^*$ ($y \leqslant y^*$).

Since $\partial_1 f_1(y^*) > 0$, the implicit function theorem yields open neighbourhoods U_1 of y^*, V_1 of $\bar{y}^* := (y_2^*, \dots, y_n^*)$ and a function $g_1 \colon V_1 \to \mathbb{R}$, such that $f_1(g_1(\bar{y}), \bar{y}) = 0$. Also, if $y \in U$ satisfies $f_1(y) = 0$, then $y_1 = g_1(\bar{y})$.

Lemma 2.1. *The function g_1 can be extended to an open set that contains $\langle \bar{x}^0, \bar{y}^0 \rangle$. Moreover, if y in $\langle x^0, y^0 \rangle$ satisfies $f_1(y) = 0$, then $y_1 = g_1(\bar{y})$.*

Proof. Recall that

$$\partial_j g_1(\bar{z}) = -\frac{\partial_j f_1(g_1(\bar{z}), \bar{z})}{\partial_1 f_1(g_1(\bar{z}), \bar{z})} \geqslant 0 \quad 2 \leqslant j \leqslant n.$$

Consider K such that

$$-\frac{\partial_j f_1(z)}{\partial_1 f_1(z)} \leqslant K \quad 2 \leqslant j \leqslant n,$$

for all z in an open neighbourhood of $\langle x^0, y^0 \rangle$ and let z in $\langle x^0, y^0 \rangle$ be such that $\bar{z} \in \overline{V_1}$ and let sequences (z_m), (w_m) in $\langle x^0, y^0 \rangle$ be such that $(\bar{z}_m)$, $(\bar{w}_m)$ are in V_1 and are both convergent to $\bar{z}$. Clearly,

$$|g_1(\bar{z}_m) - g_1(\bar{z}_p)| \leqslant K \|\bar{z}_m - \bar{z}_p\|$$

as well as

$$|g_1(\bar{z}_m) - g_1(\bar{w}_m)| \leqslant K \|\bar{z}_m - \bar{w}_m\|,$$

where euclidean norm is considered. Thus $\lim(g_1(\bar{z}_m))$ exists and its value does not depend on the sequence. We define it as $g_1(\bar{z})$. By considering a convergent subsequence of $((g_1(\bar{z}_m), \bar{z}_m))$, its limit w satisfies $\bar{w} = \bar{z}$ and one clearly has $w_1 = g_1(\bar{z})$ as well as $f_1(w) = 0$. Now, to each such w the implicit function theorem can be applied. Let us call g_2 such a function and suppose that $\bar{u}$ belongs to the domains of both g_1 and g_2. Then

$$0 = f_1(g_1(\bar{u}), \bar{u}) - f_1(g_2(\bar{u}), \bar{u}) = \partial_1 f_1(\xi, \bar{u})(g_1(\bar{u}) - g_2(\bar{u})),$$

which yields $g_1(\bar{u}) = g_2(\bar{u})$. Clearly a compactness argument combined with a maximality argument yields the conclusion. $\square$

Thus with g_1 extended to all of $\langle \bar{x}^0, \bar{y}^0 \rangle$, its graph describes the zero set of f_1 in $\langle x^0, y^0 \rangle$. Note that g_1 is isotone, so that if $\bar{z} \leqslant \bar{w}$, then $g_1(\bar{z}) \leqslant g_1(\bar{w})$. Thus for $0 \leqslant \lambda \leqslant 1$ one has

$$\begin{aligned} &f_1(\lambda g_1(\bar{z}) + (1-\lambda) g_1(\bar{w}), \lambda \bar{z} + (1-\lambda)\bar{w}) \\ &\qquad \leqslant \lambda f_1(g_1(\bar{z}), \bar{z}) + (1-\lambda) f_1(g_1(\bar{w}), \bar{w}) = 0 \\ &\qquad = f_1(g_1(\lambda \bar{z} + (1-\lambda)\bar{w}), \lambda \bar{z} + (1-\lambda)\bar{w}), \end{aligned}$$

which implies that

$$\lambda g_1(\bar{z}) + (1-\lambda) g_1(\bar{w}) \leqslant g_1(\lambda \bar{z} + (1-\lambda)\bar{w}),$$

that is, g_1 is order concave. Alternatively, from the hypotheses on F' it easily follows that g_1' is antitone, which also yields that g_1 is order concave (see [9]).

Let us now denote

$$C_1^+ := \{(y_1, \bar{y}) \colon \bar{y} \in \langle \bar{x}^0, \bar{y}^0 \rangle,\ g_1(\bar{y}) \leqslant y_1 \leqslant y_1^0\}$$

and

$$C_1^- := \{(y_1, \bar{y}) \colon \bar{y} \in \langle \bar{x}^0, \bar{y}^0 \rangle,\ x_1^0 \leqslant y_1 \leqslant g_1(\bar{y})\}.$$

The preceding remarks yield the following lemma.

Lemma 2.2. *With the above notation, it follows that*

$$C_1^+ = \{y \in \langle x^0, y^0 \rangle : f_1(y) \geqslant 0\},$$

and

$$C_1^- = \{y \in \langle x^0, y^0 \rangle : f_1(y) \leqslant 0\}.$$

Note that C_1^- is an order convex set, namely if $z, w \in C_1^-$ with $z \leqslant w$ and $0 \leqslant \lambda \leqslant 1$, then $\lambda z + (1-\lambda) w \in C_1^-$. Analogously, implicit functions can be defined on each coordinate projection of $\langle x^0, y^0\rangle$, and by considering the corresponding sets C_i^+, C_i^-, $i = 2\ldots, n$, we obtain that

$$K^+ := \{y \in \langle x^0, y^0\rangle : F(y) \geqslant 0\} = \bigcap_{i=1}^{n} C_i^+,$$

and

$$K^- := \{y \in \langle x^0, y^0\rangle : F(y) \leqslant 0\} = \bigcap_{i=1}^{n} C_i^-.$$

Now also K^- is an order convex set, while in general K^+ is not, as the following example shows.

Consider $n = 2$ and F defined by

$$\begin{aligned} f_1(y_1, y_2) &:= y_1 - y_2 - 5, \\ f_2(y_1, y_2) &:= y_1 y_2 + 6. \end{aligned} \tag{2.1}$$

If $x^0 := y^* = (3, -2)$ and $y^0 := (6, -1)$, then $F(x^0) \leqslant 0 \leqslant F(y^0)$. Note that both x^0 and y^0 are in K^+ and that $x^0 \leqslant y^0$. Also, F' is isotone and $F'(y)$ is an irreducibly diagonally dominant M-matrix whenever $y \in \langle x^0, y^0\rangle$. Clearly

$$f_2(x^0 + t(y^0 - x^0)) = (3 + 3t)(t - 2) + 6 = 3t^2 - 3t,$$

so that

$$f_2(x^0 + t(y^0 - x^0)) < 0 \quad \text{if} \quad 0 < t < 1,$$

that is, $x^0 + t(y^0 - x^0)$ does not belong to K^+ for $0 < t < 1$, i.e. K^+ is not order convex.

The following theorem, though simple, plays a fundamental role in this paper and, together with the previous example, provides a correction to the above mentioned flaw in the proof of Theorem 3.4 in [7], which consisted in implicitly assuming that K^+ is order convex.

Theorem 2.3. *Given z^0 in K^+, there exists a continuous piecewise continuously differentiable path $g\colon [0, 1] \longrightarrow K^+$ such that*

$$g(0) = z^0, \quad g(1) = y^0, \quad \text{and} \quad g(t_1) \leqslant g(t_2) \quad \text{for } t_1 \leqslant t_2.$$

As a consequence, $g'(t) \geqslant 0$, $t \in [0, 1]$, with the possible exception of a finite number of points. Such g will be called an ascending path.

Proof. The proof proceeds by mathematical induction. Consider the implicit function g_n and set $y_n^{0,0} := g_n(\bar{y}^0)$; note that here $\bar{y} := (y_1, \ldots, y_{n-1})$.

Assume first that $z_n^0 \geqslant y_n^{0,0}$. Notice that, since g_n is isotone, $g_n(\bar{y}) \leqslant y_n^{0,0}$, whence $f_n(y) \geqslant 0$ for y in $\langle x^0, y^0 \rangle$ with $y_n = z_n^0$, that is $y \in C_n^+$. Now, for $1 \leqslant i \leqslant n-1$, one has $x_i^0 \leqslant g_i \leqslant y_i^0$, so that

$$f_i(\bar{x}^0, z_n^0) \leqslant 0 \leqslant f_i(\bar{y}^0, z_n^0).$$

Thus the induction is applied to the reduced system

$$\bar{f}_i(\bar{y}) := f_i(\bar{y}, z_n^0) = 0, \quad 1 \leqslant i \leqslant n-1, \quad \bar{y} \in \langle \bar{x}^0, \bar{y}^0 \rangle.$$

Calling $s(t)$, $0 \leqslant t \leqslant 1$ an ascending path from $\bar{z}^0$ to $\bar{y}^0$ in $\langle \bar{x}^0, \bar{y}^0 \rangle$ for the reduced system, an ascending path g joining z^0 to y^0 in K^+ is obtained by sticking together the ascending path $S(t) := (s(t), z_n^0)$ with the ascending segment joining $(\bar{y}^0, z_n^0)$ to y^0.

If conversely $z_n^0 < y_n^{0,0}$, consider the segment $s(t) := \bar{z}^0 + t*(\bar{y}^0 - \bar{z}^0)$, $0 \leqslant t \leqslant 1$, and let t_0 be the infimum such that $g_n(s(t_0)) = z_n^0$. Then an ascending path joining z^0 to y^0 is obtained by ascending first along the linear segment joining z^0 to $(s(t_0), z_n^0)$, then by proceeding along $(s(t), g_n(s(t)))$, $t_0 \leqslant t \leqslant 1$, where it is employed that g_n is isotone, and finally by ascending along the segment joining $(\bar{y}^0, y_n^{0,0})$ to y^0. □

3. Brown's analytic method

The improved algorithm for Brown's method mentioned in the introduction was developed for computational purposes. In this paper, the original one is employed instead, because it allows a better description of the steps leading to the results presented here. It is not difficult to show that both algorithms generate the same Brown iterates, by taking into account that the original algorithm corresponds to the standard Gauss triangulation procedure for the linear case, while the improved algorithm corresponds to the Gauss-Jordan diagonalization procedure.

With y^0 as starting point, the next algorithm produces $B(y^0)$, the first analytic Brown iterate, now denoted y^1.

Step 1. Set $\bar{y}^0 := y^0$, $i := 1$ and $F_1(y) := (f_{1,j}(y)) := (f_j(y))$.

Step 2. Consider a first order Taylor development of $f_{i,i}$ at $\bar{y}^0$, equate it to 0 and solve for y_i, the resulting identity being $y_i = l_i(y_{i+1}, \ldots, y_n)$.

Step 3. Define the $(i+1)$st reduced system of order $n-i$ by

$$F_{i+1}(y_{i+1}, \ldots, y_n) := (f_{i+1,j}(y_{i+1}, \ldots, y_n)) = 0,$$
$$\text{where for } i+1 \leqslant j \leqslant n,$$
$$f_{i+1,j}(y_{i+1}, \ldots, y_n) := f_{i,j}(l_i(y_{i+1}, \ldots, y_n), y_{i+1}, \ldots, y_n).$$

Step 4. If $i+1 < n$, set $i := i+1$, $\bar{y}^0 := (y^0_{i+1}, \dots, y^0_n)$, and start again with step 2.

Step 5. Consider a first order Taylor development of $F_n = f_{n,n}$ centered at y^0_n, equate it to 0 and call its solution y^1_n.

Step 6. For $i = n-1$ to 1, set $y^1_i := l_i(y^1_{i+1}, \dots, y^1_n)$.

Recall that $y^* \leqslant y^1 \leqslant y^0$ and that $F(y^1) \geqslant 0$ (see [3]). This algorithm can be carried out by assuming that the Jacobian matrix F' is always nonsingular, in which case some kind of pivoting may be necessary. No pivoting has been introduced here, because in the MNT context $F'(y)$ is assumed to be a nonsingular M-matrix and this property is inherited by each F_i, $2 \leqslant i \leqslant n$, as follows by induction from the next two lemmas. As for the former, it is an adaptation of Lemma 3.3 in [5], whose proof omitted a not entirely simple argument included here.

Lemma 3.1. *For each $y \in \langle x^0, y^0 \rangle$, $F_2'(\bar{y})$ is a nonsingular* M-*matrix.*

Proof. Notice that

$$l_1(\bar{y}) = y^0_1 - \frac{1}{\partial_1 f_1(y^0)} \left[f_1(y^0) + \sum_{j=2}^{n} \partial_j f_1(y^0)(y_j - y^0_j) \right],$$

so that if $i \neq 1 \neq j$ one gets

$$\begin{aligned}\partial_j f_{2,i}(\bar{y}) &= \partial_j f_i(l_1(\bar{y}), \bar{y}) + \partial_1 f_i(l_1(\bar{y}), \bar{y}) * \partial_j l_1(\bar{y}) \\ &= \partial_j f_i(l_1(\bar{y}), \bar{y}) - \partial_1 f_i(l_1(\bar{y}), \bar{y}) * \frac{\partial_j f_1(y^0)}{\partial_1 f_1(y^0)}.\end{aligned}$$

Since F' is always an M-matrix, hence $F_2'(\bar{y})$ is a Z-matrix, i.e. its off-diagonal terms are nonpositive, because

$$\partial_j f_{2,i}(\bar{y}) \leqslant \partial_j f_i(l_1(\bar{y}), \bar{y}) \leqslant 0 \quad \text{for} \quad j \neq i.$$

With $y \in \langle \bar{x}^0, \bar{y}^0 \rangle$, consider the matrix

$$A := \begin{pmatrix} \partial_1 f_1(y^0) & \dots & \partial_n f_1(y^0) \\ \partial_1 f_2(l_1(\bar{y}), \bar{y}) & \dots & \partial_n f_2(l_1(\bar{y}), \bar{y}) \\ \vdots & \ddots & \vdots \\ \partial_1 f_n(l_1(\bar{y}), \bar{y}) & \dots & \partial_n f_n(l_1(\bar{y}), \bar{y}) \end{pmatrix}.$$

It is apparent that A is a Z-matrix and that $F'(l_1(\bar{y}), \bar{y}) \leqslant A$. Since $F'(l_1(\bar{y}), \bar{y})$ is a nonsingular M-matrix, its associated Jacobi matrix is convergent (see [12]), i.e.

$$r(D(F'(l_1(\bar{y}), \bar{y}))^{-1} * [D(F'(l_1(\bar{y}), \bar{y})) - F'(l_1(\bar{y}), \bar{y})]) < 1,$$

where r denotes the spectral radius and $D(\cdot)$ denotes the corresponding diagonal matrix. By virtue of

$$0 \leqslant D(A)^{-1} * [D(A) - A] \leqslant D(F'(l_1(\bar{y}), \bar{y}))^{-1} * [D(F'(l_1(\bar{y}), \bar{y})) - F'(l_1(\bar{y}), \bar{y})],$$

the Perron-Frobenius theory (see [12]) yields

$$r(D(A)^{-1} * [D(A) - A]) \leqslant r(D(F'(l_1(\bar{y}), \bar{y}))^{-1} * [D(F'(l_1(\bar{y}), \bar{y})) - F'(l_1(\bar{y}), \bar{y})]),$$

whence

$$r(D(A)^{-1} * [D(A) - A]) < 1.$$

Thus A itself is nonsingular and $A^{-1} \geqslant 0$ (see [12]), i.e. it is a nonsingular M-matrix. Consider now

$$M_1 := \begin{pmatrix} 1 & 0 & 0 & \dots & \dots & 0 \\ -m_{2,1} & 1 & 0 & \dots & \dots & 0 \\ -m_{3,1} & 0 & 1 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 1 & 0 \\ -m_{n,1} & 0 & \dots & \dots & 0 & 1 \end{pmatrix} \quad \text{with } m_{i,1} := \frac{\partial_1 f_i(l_1(\bar{y}), \bar{y})}{\partial_1 f_1(y^0)},$$

whence it follows that

$$M_1 * A = \begin{pmatrix} \partial_1 f_1 & \dots & \dots & \partial_n f_1 \\ 0 & & & \\ \vdots & & F_2(\bar{y})' & \\ 0 & & & \end{pmatrix}.$$

Thus $F_2'(\bar{y})$ is nonsingular, and by taking into account the above block structure, one gets

$$A^{-1} * M_1^{-1} = \begin{pmatrix} (\partial_1 f_1)^{-1} & \dots & \dots & \dots \\ 0 & & & \\ \vdots & & (F_2'(\bar{y}))^{-1} & \\ 0 & & & \end{pmatrix}.$$

Since the product on the left-hand side leaves the last $n-1$ columns in A^{-1} unchanged, one gets that $(F_2'(\bar{y}))^{-1} \geqslant 0$, which together with $F_2'(\bar{y})$ being a Z-matrix yields that $F_2'(\bar{y})$ is a nonsingular M-matrix. □

Lemma 3.2. *The following propositions hold:*
(i) F_2' *is isotone on* $\langle \bar{x}^0, \bar{y}^0 \rangle$.
(ii) $0 \leqslant F_2(\bar{y}^0)$.

Proof. (i) It easily follows by adapting Lemma 3.5 in [5].
(ii) See Theorem 3.4 (ii) in [8]. □

Let us now define, for each $1 \leqslant i \leqslant n-1$, and inductively for $j \leqslant i$

$$
\begin{aligned}
&L_{i,i}(\bar{y}) := l_i(\bar{y}), \quad \text{and} \\
&L_{i,j}(\bar{y}) := l_j(L_{i,j+1}(\bar{y}), L_{i,j+2}(\bar{y}), \dots, L_{i,i}(\bar{y}), \bar{y}) \quad \text{for } j \neq i.
\end{aligned}
$$

With this notation it is possible to give a handy representation of $f_{i,i}$ in terms of $f_{1,i} = f_i$, namely

$$f_{i,i}(\bar{y}) = f_i(\widehat{y}) \quad \text{with } \widehat{y} := (L_{i-1,1}(\bar{y}), L_{i-1,2}(\bar{y}), \dots, L_{i-1,i-1}(\bar{y}), \bar{y}).$$

Notice now that one step of Brown's method can be written in matrix form as

$$T(y^0) * (B(y^0) - y^0) = -\Phi(y^0), \tag{3.1}$$

where $\Phi(y^0) := (f_{i,i}(\bar{y}^0))$. The symbol $T(y^0)$ stands for the upper triangular matrix whose ith row is the gradient $\partial f_{i,i}(\bar{y}^0)$, that is $t_{i,k} := \partial_k f_{i,i}(\bar{y}^0)$ for $i \leqslant k$ and $t_{i,k} := 0$ for $i > k$. Notice also that $T(y^0)$ is the upper triangular part in the LU decomposition of $F'(y^0)$.

In order to analyse the variation of B in K^+, it is necessary to point out that the function Φ itself varies with the initial point y^0; this parametric dependence, which also affects T, will be considered implicit in (3.1), which is all needed here. But when considering the total differential of (3.1), one has to take into account the parametric role of y^0, as well as when considering partial derivatives of the functions involved. This is denoted here by putting the affected functions in round brackets whenever necessary. Thus, from the total differential of (3.1), one gets

$$T(y^0) * \partial B(y^0) = T(y^0) - \partial(\Phi(y^0)) - \partial(T(y^0)) * (B(y^0) - y^0). \tag{3.2}$$

Note that $\partial(T(y^0))$ is a three-linear functional.

Theorem 3.3. *The following inequalities hold componentwise:*

$$\partial(\Phi(y^0)) \leqslant T(y^0) \quad \textit{and} \quad \partial(T(y^0)) \geqslant 0.$$

Proof. The results are a consequence of the following two coupled inductive propositions for $1 \leqslant i \leqslant n-1$:

(a) For $1 \leqslant k \leqslant n$, $1 \leqslant j \leqslant i$, $i+1 \leqslant m$ the functions $L_{i,j}$ satisfy

$$\text{(a.1)} \qquad \partial_k(L_{i,j}(\bar{y}^0)) \geqslant \partial_k L_{i,j}(\bar{y}^0) = \partial_k L_{i,j} \geqslant 0,$$

and

$$\text{(a.2)} \qquad \partial_k(\partial_m L_{i,j}(\bar{y}^0)) \leqslant 0.$$

(b) For $1 \leqslant k \leqslant n$ and $j \geqslant i+1$, $f_{i+1,i+1}$ is such that

$$\text{(b.1)} \qquad \partial_k(f_{i+1,i+1}(\bar{y}^0)) \leqslant t_{i+1,k}(\bar{y}^0)$$

and

$$\text{(b.2)} \qquad \partial_k(\partial_j f_{i+1,i+1}(\bar{y}^0)) \geqslant \partial_k\partial_j f_{i+1,i+1}(\bar{y}^0) \geqslant 0.$$

Consider first the case $i=1$. As for (a.1), notice that, since

$$l_1(y^0) = y_1^0 - \frac{f_1(y^0)}{\partial_1 f_1(y^0)},$$

it follows that

$$\text{(3.3)} \qquad \partial_k(l_1(y^0)) = \delta_{1,k} - \frac{\partial_k f_1(y^0)}{\partial_1 f_1(y^0)} + f_1(y^0) * \frac{\partial^2_{k,1} f_1(y^0)}{(\partial_1 f_1(y^0))^2},$$

where δ stands for the Kronecker symbol. Thus if $k=1$,

$$\partial_1(l_1(\bar{y}^0)) = f_1(y^0) * \frac{\partial^2_{1,1} f_1(y^0)}{(\partial_1 f_1(y^0))^2} \geqslant 0 = \partial_1 l_1,$$

while if $k>1$, we have

$$\partial_k(l_1(\bar{y}^0)) = -\frac{\partial_k f_1(y^0)}{\partial_1 f_1(y^0)} + f_1(y^0) * \frac{\partial^2_{k,1} f_1(y^0)}{(\partial_1 f_1(y^0))^2} \geqslant -\frac{\partial_k f_1(y^0)}{\partial_1 f_1(y^0)} = \partial_k l_1.$$

In order to prove (a.2), recall that

$$\begin{aligned}\partial_k(\partial_m l_1(\bar{y}^0)) &= -\frac{\partial_k(\partial_m f_1(\bar{y}^0))}{\partial_1 f_1(y^0)} + \partial_m f_1(y^0) * \frac{\partial_k(\partial_1 f_1(y^0))}{(\partial_1 f_1(y^0))^2} \\ &= -\frac{\partial^2_{k,m} f_1(\bar{y}^0)}{\partial_1 f_1(y^0)} + \partial_m f_1(y^0) * \frac{\partial^2_{k,1} f_1(y^0)}{(\partial_1 f_1(y^0))^2}.\end{aligned}$$

Since $\partial^2 \geqslant 0$ and $\partial_m f_1 \leqslant 0$ because $m > 1$, (a.2) is finally obtained for $i = 1$.

To prove (b.1), using

$$\begin{aligned}\partial_k(f_{2,2}(\bar{y}^0)) &= \partial_k(f_2(l_1(\bar{y}^0), \bar{y}^0)) \\ &= \partial_1 f_2(\hat{y}^0) * \partial_k(l_1(\bar{y}^0)) + (1 - \delta_{1,k}) * \partial_k f_2(\hat{y}^0),\end{aligned}$$

and taking into account (3.3), we conclude that

$$\partial_1(f_{2,2}(\bar{y}^0)) = \partial_1 f_2(\hat{y}^0) * f_1(y^0) * \frac{\partial^2_{1,1} f_1(y^0)}{(\partial_1 f_1(y^0))^2} \leqslant 0.$$

On the other hand, if $k > 1$, then

$$\begin{aligned}\partial_k(f_{2,2}(\bar{y}^0)) &= \partial_1 f_2(\hat{y}^0) * \left(\partial_k l_1 + f_1(y^0) * \frac{\partial^2_{k,1} f_1(y^0)}{(\partial_1 f_1(y^0))^2}\right) + \partial_k f_2(\hat{y}^0) \\ &\leqslant \partial_1 f_2(\hat{y}^0) * \partial_k l_1 + \partial_k f_2(\hat{y}^0) = \partial_k f_{2,2}(\bar{y}^0).\end{aligned}$$

In order to prove now (b.2) with $i = 1$, note first that

$$\begin{aligned}\partial_k(\partial_j f_{2,2}(\bar{y}^0)) &= \partial_k(\partial_1 f_2(\hat{y}^0) * \partial_j l_1 + \partial_j f_2(\hat{y}^0)) \\ &= \partial_k(\partial_1 f_2(\hat{y}^0)) * \partial_j l_1 + \partial_k(\partial_j f_2(\hat{y}^0)) + \partial_1 f_2(\hat{y}^0) * \partial_k(\partial_j l_1).\end{aligned}$$

Since from (a.1) one has

$$\begin{aligned}\partial_k(\partial_j f_2(\hat{y}^0)) &= \partial^2_{1,j} f_2(\hat{y}^0) * \partial_k(l_1(\bar{y}^0)) + (1 - \delta_{1,k}) * \partial^2_{k,j} f_2(\hat{y}^0) \\ &\geqslant \partial^2_{1,j} f_2(\hat{y}^0) * \partial_k l_1 + (1 - \delta_{1,k}) * \partial^2_{k,j} f_2(\hat{y}^0),\end{aligned}$$

this inequality and the corresponding one for $j = 1$ yield

$$\begin{aligned}\partial_k(\partial_j f_{2,2}(\bar{y}^0)) &\geqslant [\partial^2_{1,1} f_2(\hat{y}^0) * \partial_k l_1 + (1 - \delta_{1,k}) * \partial^2_{k,1} f_2(\hat{y}^0)] * \partial_j l_1 \\ &\quad + [\partial^2_{1,j} f_2(\hat{y}^0) * \partial_k l_1 + (1 - \delta_{1,k}) * \partial^2_{k,j} f_2(\hat{y}^0)] \\ &\quad + \partial_1 f_2(\hat{y}^0) * \partial_k(\partial_j l_1).\end{aligned} \tag{3.4}$$

On the other hand,

$$\begin{aligned}\partial_k\partial_j f_{2,2}(\bar{y}^0) &= \partial_k(\partial_1 f_2(\hat{y}^0) * \partial_j l_1 + (1-\delta_{1,j}) * \partial_j f_2(\hat{y}^0))\\ &= [\partial^2_{1,1} f_2(\hat{y}^0) * \partial_k l_1 + (1-\delta_{1,k}) * \partial^2_{k,1} f_2(\hat{y}^0)] * \partial_j l_1\\ &\quad + (1-\delta_{1,j})[\partial^2_{1,j} f_2(\hat{y}^0) * \partial_k l_1 + (1-\delta_{1,k}) * \partial^2_{k,j} f_2(\hat{v}^0)].\end{aligned}$$

Since $j > 1$, going back to (3.4) one obtains

$$\partial_k(\partial_j f_{2,2}(\bar{y}^0)) \geqslant \partial_k\partial_j f_{2,2}(\bar{y}^0) + \partial_1 f_2(\hat{v}^0) * \partial_k(\partial_j l_1).$$

Taking into account (a.2) and the inequality $\partial_1 f_2 \leqslant 0$, one finally gets (b.2) for $i = 1$.

Now let us assume that (a) and (b) are valid for all $m < i \leqslant n-1$ with $i > 1$. Hence, for (a.1), it follows that

$$\begin{aligned}\partial_k(L_{i,i}(\bar{y}^0)) = \partial_k(l_i(\bar{y}^0)) &= \partial_k\left(y_i^0 - \frac{f_{i,i}(\bar{y}^0)}{\partial_i f_{i,i}(\bar{y}^0)}\right)\\ &= \delta_{i,k} - \frac{\partial_k(f_{i,i}(\bar{y}^0))}{\partial_i f_{i,i}(\bar{y}^0)} + f_{i,i}(\bar{y}^0) * \frac{\partial_k(\partial_i f_{i,i}(\bar{y}^0))}{(\partial_i f_{i,i}(\bar{y}^0))^2}\\ &\geqslant \delta_{i,k} - \frac{\partial_k(f_{i,i}(\bar{y}^0))}{\partial_i f_{i,i}(\bar{y}^0)}\\ &\geqslant \delta_{i,k} - \frac{\partial_k f_{i,i}(\bar{y}^0)}{\partial_i f_{i,i}(\bar{y}^0)}\\ &= \partial_k l_i = \partial_k L_{i,i}(\bar{y}^0) = \partial_k L_{i,i} \geqslant 0,\end{aligned}$$

where the first inequality above follows from (b.2), while the other from (b.1). This completes the first step in the inner induction from $j = i$ to $j = 1$, in order to prove (a.1). Now, since $L_{i,j}(\bar{y}^0) = l_j(\hat{y}^0)$ where $\hat{y}^0 = (L_{i,j+1}(\bar{y}^0), \ldots, L_{i,i}(\bar{y}^0), \bar{y}^0)$, one obtains

$$\begin{aligned}\partial_k(L_{i,j}(\bar{y}^0)) &= \sum_{m=j+1}^{i} \partial_m l_j * \partial_k(L_{i,m}(\bar{y}^0)) + \prod_{m=j+1}^{i}(1-\delta_{m,k}) * \partial_k l_j\\ &\geqslant \sum_{m=j+1}^{i} \partial_m l_j * \partial_k L_{i,m} + \prod_{m=j+1}^{i}(1-\delta_{m,k}) * \partial_k l_j\\ &= \partial_k L_{i,j}(\bar{y}^0) = \partial_k L_{i,j} \geqslant 0\end{aligned}$$

with $\partial_k L_{i,j} = 0$ if $k \leqslant j$. Notice that also the last inequality follows inductively from the inequality for $j = i$.

As for (a.2), another descending induction from $j = i$ to $j = 1$ is performed. Thus for $m \geqslant i + 1$, (b.2) yields

$$\begin{aligned}\partial_k(\partial_m L_{i,i}(\bar{y}^0)) &= -\frac{\partial_k(\partial_m f_{i,i}(\bar{y}^0))}{\partial_i f_{i,i}(\bar{y}^0)} + \partial_m f_{i,i}(\bar{y}^0) * \frac{\partial_k(\partial_i f_{i,i}(\bar{y}^0))}{(\partial_i f_{i,i}(\bar{y}^0))^2} \\ &\leqslant -\frac{\partial_k \partial_m f_{i,i}(\bar{y}^0)}{\partial_i f_{i,i}(\bar{y}^0)} \leqslant 0,\end{aligned}$$

while if $j < i$, the (inner) induction and (a.1) as proved for i imply

$$\begin{aligned}\partial_k(\partial_m L_{i,j}(\bar{y}^0)) &= \partial_k\left(\sum_{s=j+1}^{i} \partial_s l_j * \partial_m L_{i,s}(\bar{y}^0) + \partial_m l_j\right) \\ &= \sum_{s=j+1}^{i} \partial_k(\partial_s l_j) * \partial_m L_{i,s} \\ &\quad + \sum_{s=j+1}^{i} \partial_s l_j * \partial_k(\partial_m L_{i,s}(\bar{y}^0)) + \partial_k(\partial_m l_j) \leqslant 0.\end{aligned}$$

In order to prove (b.1), notice that from (a.1) it follows as above that

$$\begin{aligned}\partial_k(f_{i+1,i+1}(\bar{y}^0)) &= \sum_{j=1}^{i} \partial_j f_{i+1}(\hat{y}^0) * \partial_k(L_{i,j}(\bar{y}^0)) + \prod_{j=1}^{i}(1 - \delta_{k,j}) * \partial_k f_{i+1}(\hat{y}^0) \\ &\leqslant \sum_{j=1}^{i} \partial_j f_{i+1}(\hat{y}^0) * \partial_k L_{i,j} + \prod_{j=1}^{i}(1 - \delta_{k,j}) * \partial_k f_{i+1}(\hat{y}^0) \\ &= t_{i+1,k}(\bar{y}^0).\end{aligned}$$

Regarding (b.2), notice first that

$$\begin{aligned}\partial_k(\partial_j f_{i+1,i+1}(\bar{y}^0)) &= \partial_k\left(\sum_{m=1}^{i} \partial_m f_{i+1}(\hat{y}^0) * \partial_j L_{i,m}(\bar{y}^0) + \partial_j f_{i+1}(\hat{y}^0)\right) \\ &= \sum_{m=1}^{i} \partial_k(\partial_m f_{i+1}(\hat{y}^0)) * \partial_j L_{i,m} + \partial_k(\partial_j f_{i+1}(\hat{y}^0)) \\ &\quad + \sum_{m=1}^{i} \partial_m f_{i+1}(\hat{y}^0) * \partial_k(\partial_j L_{i,m}(\bar{y}^0)).\end{aligned} \tag{3.5}$$

For each term in (3.5) it is obtained from (a.1) that

$$\begin{aligned}
\partial_k(\partial_m f_{i+1}(\hat{y}^0)) &= \sum_{p=1}^{i} \partial^2_{p,m} f_{i+1}(\hat{y}^0) * \partial_k(L_{i,p}(\bar{y}^0)) + \prod_{p=1}^{i}(1-\delta_{k,p}) * \partial^2_{k,m} f_{i+1}(\hat{y}^0)\\
&\geqslant \sum_{p=1}^{i} \partial^2_{p,m} f_{i+1}(\hat{y}^0) * \partial_k L_{i,p} + \prod_{p=1}^{i}(1-\delta_{k,p}) * \partial^2_{k,m} f_{i+1}(\hat{y}^0).
\end{aligned}$$

The inequalities in (a) now imply (b.2) as in the case $i = 1$. □

Theorem 3.4. *If $z^0 \in K^+$, then the Brown iterates satisfy $z^k \leqslant y^k$, $k = 1, 2, \ldots$.*

Proof. Theorem 3.3 and (3.2) imply that $\partial B \geqslant 0$ in K^+, because T is always a nonsingular M-matrix. By considering an ascending path g in K^+ joining z^0 to y^0 as in Theorem 2.3, it follows that

$$y^1 - z^1 = B(y^0) - B(z^0) = \int_0^1 \partial_t(B(g(t))\,\mathrm{d}t = \int_0^1 \partial B(g(t)) * \partial g(t)\,\mathrm{d}t \geqslant 0.$$

A simple induction completes the proof. □

4. The analytic Fourier iterations

The Fourier iterations for Brown's method have been introduced in [3] in analogy with the Newton-Fourier iterations. These Brown-Fourier iterates give us a monotone bracketing of y^*, which is moreover contained in the corresponding Newton-Fourier bracketing (see [3]). Their description in the framework of the original Brown algorithm now follows.

Step 1′. Set $\bar{x}^0 := x^0$ $(i = 1)$ and $F_1^-(x) = (f_{1,j}^-(x)) := (f_j(x)) = F(x)$.

Step 2′. Consider the affine approximation of $f_{i,i}^-$ centered at $\bar{x}^0$ with the gradient values of $f_{i,i}$ at $\bar{y}^0$ and solve for x_i, i.e. $x_i = l_i^-(x_{i+1}, \ldots, x_n)$.

Step 3′. Define the $(i+1)$th reduced lower system

$$\begin{aligned}
F_{i+1}^-(x_{i+1}, \ldots, x_n) &:= (f_{i+1,j}^-(x_{i+1}, \ldots, x_n)) = 0 \quad \text{where}\\
f_{i+1,j}^-(x_{i+1}, \ldots, x_n) &:= f_{i,j}^-(l_i^-(x_{i+1}, \ldots, x_n), x_{i+1}, \ldots, x_n) \quad \text{for } i+1 \leqslant j \leqslant n.
\end{aligned}$$

Step 4′. If $i + 1 < n$, set $i := i + 1$, $\bar{x}^0 := (x_{i+1}^0, \ldots, x_n^0)$ and start again with step 2'.

Step 5′. Consider a first order approximation of $f_{n,n}^-$ at x_n^0 with the slope given by $\partial_n f_{n,n}(\bar{y}^0)$, equate it to 0 and call its solution x_n^1.

Step 6′. For $i = n-1$ to 1 define $x_i^1 := l_i^-(x_{i+1}^1, \dots, x_n^1)$.

The x^1 thus obtained is the first Fourier iterate corresponding to the Brown iterate y^1. Recall that $x^0 \leqslant x^1 \leqslant y^*$ and $F(x^1) \leqslant 0$ (see [3]).

Analogously to the previous section, let us now define inductively, for $1 \leqslant i \leqslant n-1$ and for $j \leqslant i$,

$$L_{i,i}^-(\bar{x}) := l_i^-(\bar{x}) = x_i^0 - \frac{1}{\partial_1 f_{i,i}(\bar{y}^0)} \left[f_{i,i}^-(\bar{x}^0) + \sum_{j=i+1}^{n} \partial_j f_{i,i}(\bar{y}^0)(x_j - x_j^0) \right],$$

and for $j \neq i$,

$$L_{i,j}^-(\bar{x}) := l_j^-(L_{i,j+1}^-(\bar{x}), L_{i,j+2}^-(\bar{x}), \dots, L_{i,i}^-(\bar{x}), \bar{x}).$$

Here one easily gets

$$f_{i,i}^-(\bar{x}) = f_i(\widehat{x}) \quad \text{with } \widehat{x} := (L_{i-1,1}^-(\bar{x}), L_{i-1,2}^-(\bar{x}), \dots, L_{i-1,i-1}^-(\bar{x}), \bar{x}).$$

Lemma 4.1. *$f_{i,i}^-(\bar{x}^0) \leqslant 0$ for $2 \leqslant i \leqslant n$.*

Proof. Consider first $i = 2$. Then

$$\begin{aligned} f_{2,2}^-(\bar{x}^0) &= f_2(\widehat{x}^0) = f_2(L_{1,1}^-(\bar{x}^0), \bar{x}^0) = f_2(l_1^-(\bar{x}^0), \bar{x}^0) - f_2(x^0) + f_2(x^0) \\ &\leqslant \partial_1 f_2(l_1^-(\bar{x}^0), \bar{x}^0) * (l_1^-(\bar{x}^0) - x_1^0) + f_2(x^0) \\ &= \partial_1 f_2(l_1^-(\bar{x}^0), \bar{x}^0) * \left(-\frac{f_1(x^0)}{\partial_1 f_1(y^0)} \right) + f_2(x^0) \\ &\leqslant f_2(x^0) \leqslant 0, \end{aligned}$$

where the first inequality is a consequence of the order convexity. The proof is now completed by induction. Assume that for some i, $2 \leqslant i \leqslant n$,

$$f_{j,j}^-(\bar{x}^0) \leqslant 0 \quad \text{if} \quad 2 \leqslant j \leqslant i-1 < n.$$

Notice first that, because of the order convexity,

$$\begin{aligned} f_{i,i}^-(\bar{x}^0) &= f_i(\widehat{x}^0) = f_i(\widehat{x}^0) - f_i(x^0) + f_i(x^0) \\ &\leqslant \sum_{j=1}^{i-1} \partial_j f_i(\widehat{x}^0) * (L_{i-1,j}^-(\bar{x}^0) - x_j^0) + f_i(x^0). \end{aligned} \tag{4.1}$$

Now it is necessary to prove that

$$L_{i-1,j}^-(\bar{x}^0) - x_j^0 \geqslant 0 \quad \text{for} \quad 1 \leqslant j \leqslant i-1. \tag{4.2}$$

This proof is done by means of an (inner) induction from $i-1$ to 1. Notice first that

$$L^-_{i-1,i-1}(\bar{x}^0) - x^0_{i-1} = -\frac{f^-_{i-1,i-1}(\bar{x}^0)}{\partial_{i-1} f_{i-1,i-1}(\bar{y}^0)} \geqslant 0.$$

For the general term, one has

$$\begin{aligned} L^-_{i-1,j}(\bar{x}^0) - x^0_j &= l_j(\widehat{x}^0) - x^0_j \\ &= -\frac{1}{\partial_j f_{j,j}(\bar{y}^0)} * \left[f^-_{j,j}(\bar{x}^0) + \sum_{k=j+1}^{i-1} \partial_k f_{j,j}(\bar{y}^0)(L^-_{i-1,k}(\bar{x}^0) - x^0_k) \right], \end{aligned}$$

which makes it clear that the inner and outer inductions imply (4.2). By applying these inequalities in (4.1), one finally gets

$$f^-_{i,i}(\bar{x}^0) \leqslant f_i(\bar{x}^0) \leqslant 0,$$

and the proof is complete. □

Denoting $\Phi^-(x^0) := (f^-_{i,i}(\bar{x}^0))$ and $B(x^0, y^0) := x^1$, for the Fourier iterate one has the system

$$T(y^0) * (B(x^0, y^0) - x^0) = -\Phi^-(x^0). \tag{4.3}$$

Let us assume first that x^0 is kept fixed while y^0 is allowed to vary. The corresponding total differential of (4.3) then yields

$$T(y^0) * \partial B(x^0, y^0) = -\partial(\Phi^-(x^0)) - \partial(T(y^0)) * (B(x^0, y^0) - x^0). \tag{4.4}$$

Lemma 4.2. *With the notation as in* (4.4), *it follows that* $\partial(\Phi^-(x^0)) \geqslant 0$.

Proof. Note first that $\partial_k(f^-_{1,1}(x^0)) = \partial_k(f_1(x^0)) = 0$, for $1 \leqslant k \leqslant n$.

Since $f^-_{i,i}(\bar{x}^0) \leqslant 0$, $1 \leqslant i \leqslant n$, and by virtue of (b.2) in Theorem 3.3, it is also clear that

$$\begin{aligned} \partial_k(L^-_{i,i}(\bar{x}^0)) = \partial_k(l^-_i(\bar{x}^0)) &= \partial_k\left(x^0_i - \frac{f^-_{i,i}(\bar{x}^0)}{\partial_i f_{i,i}(\bar{y}^0)} \right) \\ &= f^-_{i,i}(\bar{x}^0) * \frac{\partial_k(\partial_i f_{i,i}(\bar{y}^0))}{(\partial_i f_{i,i}(\bar{y}^0))^2} \leqslant 0. \end{aligned}$$

It now inductively follows, from $j = i$ to $j = 1$, that

$$\begin{aligned} \partial_k(L^-_{i,j}(\bar{x}^0)) &= \sum_{m=j+1}^{i} \partial_m l^-_j * \partial_k(L^-_{i,m}(\bar{x}^0)) \\ &= -\sum_{m=j+1}^{i} \frac{\partial_m f_{j,j}(\bar{y}^0)}{\partial_j f_{j,j}(\bar{y}^0)} * \partial_k(L^-_{i,m}(\bar{x}^0)) \leqslant 0. \end{aligned}$$

This inequality yields

$$\partial_k(f^-_{i+1,i+1}(\overline{x}^0)) = \partial_k(f_{i+1}(\widehat{x}^0)) = \sum_{j=1}^{i} \partial_j f_{i+1}(\widehat{x}^0) * \partial_k(L^-_{i,j}(\overline{x}^0)) \geqslant 0,$$

because $\partial_j f_{i+1} \leqslant 0$ for $j \leqslant i$, which completes the proof. $\square$

Theorem 4.3. *If $z^0 \in K^+$, then $B(x^0, y^0) \leqslant B(x^0, z^0)$.*

Proof. Lemma 4.2 and Theorem 3.3 applied in (4.4) imply that $\partial B(x^0, y) \leqslant 0$ for y in K^+. The conclusion then follows as in Theorem 3.4. $\square$

Suppose now that y^0 is held fixed while x^0 varies. In this case, from the corresponding differential in (4.3), it is easy to see that

$$T(y^0) * \partial B(x^0, y^0) = T(y^0) - \partial(\Phi^-(x^0)). \tag{4.5}$$

Lemma 4.4. *With the notation as in* (4.5) *it follows that $\partial(\Phi^-(x^0)) \leqslant T(y^0)$.*

Proof. Clearly one has that

$$\partial_k(f_1(x^0)) = \partial_k f_1(x^0) \leqslant \partial_k f_1(y^0) = t_{1,k}(y^0).$$

Assuming now for some $i < n-1$ and all k, $1 \leqslant k \leqslant n$, that

$$\partial_k(f^-_{i,i}(\overline{x}^0)) \leqslant t_{i,k}(y^0) = \partial_k f_{i,i}(y^0) = t_{i,k}(y^0),$$

we obtain that

$$\begin{aligned}
\partial_k(l^-_i(\overline{x}^0)) = \partial_k(L^-_{i,i}(\overline{x}^0)) &= \partial_k\left(x^0_i - \frac{f^-_{i,i}(\overline{x}^0)}{\partial_i f_{i,i}(\overline{y}^0)}\right) \\
&= \delta_{i,k} - \frac{\partial_k(f^-_{i,i}(\overline{x}^0))}{\partial_i f_{i,i}(\overline{y}^0)} \\
&\geqslant \partial_k l^-_i = \partial_k l_i \geqslant 0,
\end{aligned}$$

so that

$$\begin{aligned}
\partial_k(L^-_{i,j}(\overline{x}^0)) &= \partial_k(l^-_j(\widehat{x}^0)) \\
&= \sum_{m=j+1}^{i} \partial_m l^-_j * \partial_k(L^-_{i,m}(\overline{x}^0)) + \prod_{m=j+1}^{i} (1-\delta_{m,k}) * \partial_k(l^-_j) \\
&\geqslant \sum_{m=j+1}^{i} \partial_m l^-_j * \partial_k L^-_{i,m} + \prod_{m=j+1}^{i} (1-\delta_{m,k}) * \partial_k l^-_j \\
&= \sum_{m=j+1}^{i} \partial_m l_j * \partial_k L_{i,m} + \prod_{m=j+1}^{i} (1-\delta_{m,k}) * \partial_k l_j \\
&= \partial_k L^-_{i,j}(\overline{y}^0) = \partial_k L_{i,j} \geqslant 0.
\end{aligned}$$

Since one has

$$\partial_m f_{i+1}(\widehat{x}^0) \leqslant \partial_m f_{i+1}(\widehat{y}^0) \leqslant 0,$$

for $1 \leqslant m \leqslant i$, one obtains

$$\begin{aligned}\partial_k(f^-_{i+1,i+1}(\bar{x}^0)) &= \partial_k(f_{i+1}(\widehat{x}^0)) \\ &= \sum_{m=1}^{i} \partial_m f_{i+1}(\widehat{x}^0) * \partial_k(L^-_{i,m}(\bar{x}^0)) + \prod_{m=1}^{i} (1-\delta_{k,m}) * \partial_k f_{i+1}(\widehat{x}^0) \\ &\leqslant \sum_{m=1}^{i} \partial_m f_{i+1}(\widehat{y}^0) * \partial_k L_{i,m} + \prod_{m=1}^{i} (1-\delta_{k,m}) * \partial_k f_{i+1}(\widehat{y}^0) \\ &= t_{i+1,k}(y^0),\end{aligned}$$

which completes the proof. □

Theorem 4.5. *If $w^0 \in K^-$, then $B(x^0, y^0) \leqslant B(w^0, y^0)$.*

Proof. Lemma 4.4, when applied in (4.5), implies that $\partial(B(x, y^0)) \geqslant 0$ for x in K^-. Since K^- is an order convex set, the ascending segment joining x^0 to w^0 is contained in K^-, so that the argument in the proofs of Theorems 3.3 and 4.3 applies again and yields the conclusion. □

Corollary 4.6. *If $w^0 \in K^-$ and $z^0 \in K^+$, then $B(x^0, y^0) \leqslant B(w^0, z^0)$.*

5. An example

An illustration is briefly discussed in this section. It deals with Chandrasekhar's equation and suggests that, given the quadratic convergence of Brown's analytic method, one should only expect a modest improvement in the number of iterations by conveniently choosing the starting (upper) point.

Consider thus Chandrasekhar's equation, namely

$$v(t) = 1 - \frac{1}{4}\int_0^1 \left(\frac{t}{s+t} * \frac{1}{v(s)}\right) \mathrm{d}s, \quad 0 \leqslant t \leqslant 1.$$

The approach to dealing with this equation follows [4] and [8]. For $h := \frac{1}{64}$, the trapezoidal integration rule is applied at the points ih, $0 \leqslant i \leqslant 64$. Taking into account that $v(0) = 1$, the resulting nonlinear system is

$$F(x) := (f_i(x)) = 0, \quad 1 \leqslant i \leqslant 64,$$

where

$$f_i(x) := x_i + \frac{1}{4}\left[w_0 + \sum_{j=1}^{64} w_j * \frac{i}{i+j} * \frac{1}{x_j}\right] - 1,$$

$$w_0 := w_{64} := \frac{h}{2} \quad w_j := h, \quad 1 \leqslant j \leqslant 63.$$

Consider two different starting upper points y^0 whose coordinates are all equal to 1 in one case and to 5 in the other. It is easily verified that $F(y^0) \geqslant 0$ in both cases. Consider also the corresponding Fourier iterations with $x_i^0 := .5$, $1 \leqslant i \leqslant 64$, for which it is also easy to see that $F(x^0) \leqslant 0$. The stopping criteria for Brown and Fourier iterations are given by the first k for which the function residues satisfy, respectively,

$$\|F(y^k)\|_\infty < \varepsilon := .5 * 10^{-13}, \quad \text{and} \quad \|F(x^k)\|_\infty < \varepsilon.$$

The computations have been carried out with the double precision of Fortran 77. The table shows the values of the iterates approximating $v(1)$, namely y_{64}^k, as well as the values of x_{64}^k. The exact digits are underlined. The final k in each column of values is the one for which the function values satisfy the corresponding stopping criterion given above.

k	y_{64}^k	x_{64}^k	y_{64}^k	x_{64}^k
0	5.	.5	1.	.5
1	.808462758084	.$\underline{7}$89714505200	.$\underline{799}$636685607	.$\underline{79}$3434227609
2	.$\underline{799}$218390107	.$\underline{7991}$26316604	.$\underline{799194}$762887	.$\underline{7991}$84364766
3	.$\underline{79919470}$2734	.$\underline{799194700}$358	.$\underline{799194702574}$	.$\underline{799194702}$544
4	.$\underline{799194702574}$	.$\underline{799194702574}$		.$\underline{799194702574}$

It is worth pointing out not only the consistency of the table with the results in the paper, but also that Brown iterates appear to converge faster than the corresponding Fourier iterates. This is a well established fact by Ostrowski when $n = 1$, namely for one dimensional analytic Newton-Fourier iterations (see [10]).

Final comment. *The extension of the results in the paper to discretized Brown iterations will be analyzed elsewhere, along with their possible application to the comparison of discretized Brown and Newton iterations.*

Acknowledgement. The referee's constructive criticism has been very helpful for the improvement of the manuscript.

References

[1] *R. Brent*: Some efficient algorithms for solving systems of nonlinear equations. SIAM J. Numer. Anal. *10* (1973), 327–344.

[2] *K. Brown*: A quadratically convergent Newton-like method based upon Gaussian elimination. SIAM J. Numer. Anal. *6* (1969), 560–569.

[3] *A. Frommer*: Monotonicity of Brown's method. Z. Angew. Math. Mech. *68* (1988), 101–109.

[4] *A. Frommer*: Comparison of Brown's and Newton's method in the monotone case. Numer. Math. *52* (1988), 511–521.

[5] *J. P. Milaszewicz, S. Abdel Masih*: Elimination and fixed point iterations. Comput. Math. Appl. *25* (1993), 43–53.

[6] *J. P. Milaszewicz*: Comparison theorems for monotone Newton-Fourier iterations and applications in functional elimination. Linear Algebra Appl. *220* (1995), 343–357.

[7] *J. P. Milaszewicz*: Comparison theorems for a third order method. Appl. Math. Lett. *10* (1997), 17–21.

[8] *J. P. Milaszewicz*: On Brown's and Newton's methods with convexity hypotheses. J. Comput. Appl. Math. *150* (2002), 1–24.

[9] *J. M. Ortega, W. C. Rheinboldt*: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York-London, 1970.

[10] *A. M. Ostrowski*: Solution of Equations and Systems of Equations. Academic Press, New York-London, 1960.

[11] *H. Schwetlick*: Numerische Lösung Nichtlinearer Gleichungssysteme. R. Oldenburg Verlag, München, 1979.

[12] *R. S. Varga*: Matrix Iterative Analysis. Prentice-Hall, Englewood Cliffs, 1962.

Author's address: *J. P. Milaszewicz*, Instituto Argentino de Matemática, CONICET, Saavedra 15, 1083 Buenos Aires, Argentina, e-mail: `milas@mate.dm.uba.ar`.

49 (2004) APPLICATIONS OF MATHEMATICS No. 2, 165–184

ON BROWN'S METHOD WITH CONVEXITY HYPOTHESES

Juan Pedro Milaszewicz, Buenos Aires

(Received April 29, 2002, in revised version September 6, 2003)

Abstract. Given two initial points generating monotone convergent Brown iterations in the context of the monotone Newton theorem (MNT), it is proved that if one of them is an upper bound of the other, then the same holds for each pair of respective terms in the Brown sequences they generate. This comparison result is carried over to the corresponding Brown-Fourier iterations. An illustration is discussed.

Keywords: nonlinear systems, convex functions, Brown's method, monotone convergence, Fourier iterates

MSC 2000: 65H10

1. Introduction

For a continuously differentiable function $F\colon D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$, consider the associated system

$$F(x) = 0. \tag{1.1}$$

In order to find an approximate solution of (1.1) when $n > 1$, Brown proposed a combination of the one dimensional Newton method with a Gauss-Seidel-like extension to the nonlinear case of the Gaussian elimination process. Quadratic convergence for Brown's analytic method was established (see [2]), but the original algorithm had a cost of $O(n^4)$ algebraic operations per iteration, which made it unattractive when compared with Newton's method. Subsequently, this figure was reduced to $O(n^3)$ with a better implementation (see [1] and [11]). With this new algorithm and in the context of the MNT (see [9]), Frommer proved a monotone Brown theorem (see [3]), as well as that Brown's analytic method converges componentwise at least as fast as Newton's analytic method (see [4]); these results include the corresponding ones for the respective Fourier iterations. Comparison theorems hold for Newton's method (see [6]) in the MNT framework. Thus it is interesting and can be useful to obtain analogous results for Brown's method in the same context, and this is the aim of the present paper. A typical comparison result proved here in that framework is that, if $z^0 \leqslant y^0$ are such that $0 \leqslant F(y^0)$, $0 \leqslant F(z^0)$, and if $F(y^*) = 0$, then $y^* \leqslant z^k \leqslant y^k$ for all Brown iterates with starting values z^0 and y^0, respectively, where the inequalities are understood componentwise.

It is convenient to recall that the MNT context arises naturally when mildly nonlinear problems and certain nonlinear integral equations are discretized. For instance, when discretizing $\Delta u = \mathrm{e}^u$ by means of second central difference quotients (see [9]), and when discretizing Chandrasekhar's integral equation (see the last section). Also, if $\Delta u = \mathrm{e}^u$ is transformed, via Green's function, into an equivalent integral equation, then its discretization by means of a trapezoidal rule generates the same context again.

Although the comparison statements for Brown's method are formally similar to those for Newton's method (see [6]), they require proofs that are much more involved. In order to simplify cumbersome calculations, the main idea here consists in proving that the Brown function assigning to each initial point its first Brown iterate, is a monotone function with respect to the coordinatewise standard order on a convenient domain. This idea has also been employed for a third order method in [7] and, though the general approach in both papers is similar, the proofs regarding the two methods are essentially unrelated.

The comparison results for Brown's method are also extended to the corresponding Fourier iterations which coupled with the Brown iterations provide bracketings, i.e. lower and upper bounds of the solution of (1.1).

The outline of the paper is as follows. The next section describes the framework and some significant conclusions are drawn regarding the geometry of the zero manifolds of f_i, $1 \leqslant i \leqslant n$; the section contains the only common features with [7], while correcting a minor flaw in a proof of that paper. The third section contains the main results for Brown's analytic method, while in the fourth their extensions to the Brown-Fourier iterations are proved as well. In the fifth section, a numerical example illustrates the results.

2. Definitions and basic results

It is assumed in the sequel that $x^0 < y^0$, i.e. $x_i^0 < y_i^0$, $1 \leqslant i \leqslant n$, and that

$$\langle x^0, y^0 \rangle := \{x\colon\ x^0 \leqslant x \leqslant y^0\} \subset D \quad \text{and} \quad F(x^0) \leqslant 0 \leqslant F(y^0).$$

Also for y in D, $F'(y)$ is a nonsingular M-matrix (see [12]), i.e.

$$(F'(y))_{i,j} \leqslant 0 \quad \text{for} \quad i \neq j \quad \text{and} \quad F'(y)^{-1} \geqslant 0.$$

It is also assumed that the Jacobian matrix F' is isotone, that is

$$F'(x) \leqslant F'(y) \quad \text{if} \quad x \leqslant y.$$

Recall that if F' is isotone, then F is order convex, namely

$$F(\lambda x + (1-\lambda)y) \leqslant \lambda F(x) + (1-\lambda)F(y), \quad \text{if} \quad x \leqslant y, \quad \text{and} \quad \lambda \in (0,1).$$

The MNT ensures that there exists a unique y^* in $\langle x^0, y^0 \rangle$, for which (1.1) holds (see Chapter 13 in [9]). Moreover, the Newton iterates with starting point y^0 converge decreasingly, while their Fourier counterparts with starting point x^0 do so increasingly, and the bracketing they both determine converges quadratically to 0. As a consequence, if y in $\langle x^0, y^0 \rangle$ satisfies $F(y) \geqslant 0$ ($F(y) \leqslant 0$), then $y \geqslant y^*$ ($y \leqslant y^*$).

Since $\partial_1 f_1(y^*) > 0$, the implicit function theorem yields open neighbourhoods U_1 of y^*, V_1 of $\overline{y}^* := (y_2^*, \dots, y_n^*)$ and a function $g_1\colon\ V_1 \to \mathbb{R}$, such that $f_1(g_1(\overline{y}), \overline{y}) = 0$. Also, if $y \in U$ satisfies $f_1(y) = 0$, then $y_1 = g_1(\overline{y})$.

Lemma 2.1. *The function g_1 can be extended to an open set that contains $\langle \overline{x}^0, \overline{y}^0 \rangle$. Moreover, if y in $\langle x^0, y^0 \rangle$ satisfies $f_1(y) = 0$, then $y_1 = g_1(\overline{y})$.*

Proof. Recall that

$$\partial_j g_1(\overline{z}) = -\frac{\partial_j f_1(g_1(\overline{z}), \overline{z})}{\partial_1 f_1(g_1(\overline{z}), \overline{z})} \geqslant 0 \quad 2 \leqslant j \leqslant n.$$

Consider K such that

$$-\frac{\partial_j f_1(z)}{\partial_1 f_1(z)} \leqslant K \quad 2 \leqslant j \leqslant n,$$

for all z in an open neighbourhood of $\langle x^0, y^0 \rangle$ and let z in $\langle x^0, y^0 \rangle$ be such that $\overline{z} \in \overline{V_1}$ and let sequences (z_m), (w_m) in $\langle x^0, y^0 \rangle$ be such that $(\overline{z}_m)$, $(\overline{w}_m)$ are in V_1 and are both convergent to $\overline{z}$. Clearly,

$$|g_1(\overline{z}_m) - g_1(\overline{z}_p)| \leqslant K \|\overline{z}_m - \overline{z}_p\|$$

as well as

$$|g_1(\overline{z}_m) - g_1(\overline{w}_m)| \leqslant K \|\overline{z}_m - \overline{w}_m\|,$$

where euclidean norm is considered. Thus $\lim(g_1(\overline{z}_m))$ exists and its value does not depend on the sequence. We define it as $g_1(\overline{z})$. By considering a convergent subsequence of $((g_1(\overline{z}_m), \overline{z}_m))$, its limit w satisfies $\overline{w} = \overline{z}$ and one clearly has $w_1 = g_1(\overline{z})$ as well as $f_1(w) = 0$. Now, to each such w the implicit function theorem can be applied. Let us call g_2 such a function and suppose that $\overline{u}$ belongs to the domains of both g_1 and g_2. Then

$$0 = f_1(g_1(\overline{u}), \overline{u}) - f_1(g_2(\overline{u}), \overline{u}) = \partial_1 f_1(\xi, \overline{u})(g_1(\overline{u}) - g_2(\overline{u})),$$

which yields $g_1(\overline{u}) = g_2(\overline{u})$. Clearly a compactness argument combined with a maximality argument yields the conclusion. $\square$

Thus with g_1 extended to all of $\langle \overline{x}^0, \overline{y}^0 \rangle$, its graph describes the zero set of f_1 in $\langle x^0, y^0 \rangle$. Note that g_1 is isotone, so that if $\overline{z} \leqslant \overline{w}$, then $g_1(\overline{z}) \leqslant g_1(\overline{w})$. Thus for $0 \leqslant \lambda \leqslant 1$ one has

$$\begin{aligned} &f_1(\lambda g_1(\overline{z}) + (1-\lambda) g_1(\overline{w}), \lambda \overline{z} + (1-\lambda)\overline{w}) \\ &\qquad \leqslant \lambda f_1(g_1(\overline{z}), \overline{z}) + (1-\lambda) f_1(g_1(\overline{w}), \overline{w}) = 0 \\ &\qquad = f_1(g_1(\lambda \overline{z} + (1-\lambda)\overline{w}), \lambda \overline{z} + (1-\lambda)\overline{w}), \end{aligned}$$

which implies that

$$\lambda g_1(\overline{z}) + (1-\lambda) g_1(\overline{w}) \leqslant g_1(\lambda \overline{z} + (1-\lambda)\overline{w}),$$

that is, g_1 is order concave. Alternatively, from the hypotheses on F' it easily follows that g_1' is antitone, which also yields that g_1 is order concave (see [9]).

Let us now denote

$$C_1^+ := \{(y_1, \overline{y}) \colon\ \overline{y} \in \langle \overline{x}^0, \overline{y}^0 \rangle,\ g_1(\overline{y}) \leqslant y_1 \leqslant y_1^0\}$$

and

$$C_1^- := \{(y_1, \overline{y}) \colon\ \overline{y} \in \langle \overline{x}^0, \overline{y}^0 \rangle,\ x_1^0 \leqslant y_1 \leqslant g_1(\overline{y})\}.$$

The preceding remarks yield the following lemma.

Lemma 2.2. *With the above notation, it follows that*

$$C_1^+ = \{y \in \langle x^0, y^0 \rangle : f_1(y) \geqslant 0\},$$

and

$$C_1^- = \{y \in \langle x^0, y^0 \rangle : f_1(y) \leqslant 0\}.$$

Note that C_1^- is an order convex set, namely if $z, w \in C_1^-$ with $z \leqslant w$ and $0 \leqslant \lambda \leqslant 1$, then $\lambda z + (1-\lambda) w \in C_1^-$. Analogously, implicit functions can be defined on each coordinate projection of $\langle x^0, y^0 \rangle$, and by considering the corresponding sets C_i^+, C_i^-, $i = 2 \ldots, n$, we obtain that

$$K^+ := \{y \in \langle x^0, y^0 \rangle : \; F(y) \geqslant 0\} = \bigcap_{i=1}^{n} C_i^+,$$

and

$$K^- := \{y \in \langle x^0, y^0 \rangle : \; F(y) \leqslant 0\} = \bigcap_{i=1}^{n} C_i^-.$$

Now also K^- is an order convex set, while in general K^+ is not, as the following example shows.

Consider $n = 2$ and F defined by

$$\begin{aligned} f_1(y_1, y_2) &:= y_1 - y_2 - 5, \\ f_2(y_1, y_2) &:= y_1 y_2 + 6. \end{aligned} \tag{2.1}$$

If $x^0 := y^* = (3, -2)$ and $y^0 := (6, -1)$, then $F(x^0) \leqslant 0 \leqslant F(y^0)$. Note that both x^0 and y^0 are in K^+ and that $x^0 \leqslant y^0$. Also, F' is isotone and $F'(y)$ is an irreducibly diagonally dominant M-matrix whenever $y \in \langle x^0, y^0 \rangle$. Clearly

$$f_2(x^0 + t(y^0 - x^0)) = (3 + 3t)(t - 2) + 6 = 3t^2 - 3t,$$

so that

$$f_2(x^0 + t(y^0 - x^0)) < 0 \quad \text{if} \quad 0 < t < 1,$$

that is, $x^0 + t(y^0 - x^0)$ does not belong to K^+ for $0 < t < 1$, i.e. K^+ is not order convex.

The following theorem, though simple, plays a fundamental role in this paper and, together with the previous example, provides a correction to the above mentioned flaw in the proof of Theorem 3.4 in [7], which consisted in implicitly assuming that K^+ is order convex.

Theorem 2.3. *Given z^0 in K^+, there exists a continuous piecewise continuously differentiable path $g\colon [0,1] \longrightarrow K^+$ such that*

$$g(0) = z^0, \quad g(1) = y^0, \quad \textit{and} \quad g(t_1) \leqslant g(t_2) \quad \textit{for } t_1 \leqslant t_2.$$

As a consequence, $g'(t) \geqslant 0$, $t \in [0,1]$, with the possible exception of a finite number of points. Such g will be called an ascending path.

P r o o f. The proof proceeds by mathematical induction. Consider the implicit function g_n and set $y_n^{0,0} := g_n(\bar{y}^0)$; note that here $\bar{y} := (y_1, \ldots, y_{n-1})$.

Assume first that $z_n^0 \geqslant y_n^{0,0}$. Notice that, since g_n is isotone, $g_n(\bar{y}) \leqslant y_n^{0,0}$, whence $f_n(y) \geqslant 0$ for y in $\langle x^0, y^0 \rangle$ with $y_n = z_n^0$, that is $y \in C_n^+$. Now, for $1 \leqslant i \leqslant n-1$, one has $x_i^0 \leqslant g_i \leqslant y_i^0$, so that

$$f_i(\bar{x}^0, z_n^0) \leqslant 0 \leqslant f_i(\bar{y}^0, z_n^0).$$

Thus the induction is applied to the reduced system

$$\overline{f}_i(\bar{y}) := f_i(\bar{y}, z_n^0) = 0, \quad 1 \leqslant i \leqslant n-1, \quad \bar{y} \in \langle \bar{x}^0, \bar{y}^0 \rangle .$$

Calling $s(t)$, $0 \leqslant t \leqslant 1$ an ascending path from $\bar{z}^0$ to $\bar{y}^0$ in $\langle \bar{x}^0, \bar{y}^0 \rangle$ for the reduced system, an ascending path g joining z^0 to y^0 in K^+ is obtained by sticking together the ascending path $S(t) := (s(t), z_n^0)$ with the ascending segment joining $(\bar{y}^0, z_n^0)$ to y^0.

If conversely $z_n^0 < y_n^{0,0}$, consider the segment $s(t) := \bar{z}^0 + t*(\bar{y}^0 - \bar{z}^0)$, $0 \leqslant t \leqslant 1$, and let t_0 be the infimum such that $g_n(s(t_0)) = z_n^0$. Then an ascending path joining z^0 to y^0 is obtained by ascending first along the linear segment joining z^0 to $(s(t_0), z_n^0)$, then by proceeding along $(s(t), g_n(s(t)))$, $t_0 \leqslant t \leqslant 1$, where it is employed that g_n is isotone, and finally by ascending along the segment joining $(\bar{y}^0, y_n^{0,0})$ to y^0. □

3. Brown's analytic method

The improved algorithm for Brown's method mentioned in the introduction was developed for computational purposes. In this paper, the original one is employed instead, because it allows a better description of the steps leading to the results presented here. It is not difficult to show that both algorithms generate the same Brown iterates, by taking into account that the original algorithm corresponds to the standard Gauss triangulation procedure for the linear case, while the improved algorithm corresponds to the Gauss-Jordan diagonalization procedure.

With y^0 as starting point, the next algorithm produces $B(y^0)$, the first analytic Brown iterate, now denoted y^1.

Step 1. Set $\bar{y}^0 := y^0$, $i := 1$ and $F_1(y) := (f_{1,j}(y)) := (f_j(y))$.

Step 2. Consider a first order Taylor development of $f_{i,i}$ at $\bar{y}^0$, equate it to 0 and solve for y_i, the resulting identity being $y_i = l_i(y_{i+1}, \ldots, y_n)$.

Step 3. Define the $(i+1)$st reduced system of order $n-i$ by

$$F_{i+1}(y_{i+1}, \ldots, y_n) := (f_{i+1,j}(y_{i+1}, \ldots, y_n)) = 0,$$
$$\text{where for } i+1 \leqslant j \leqslant n,$$
$$f_{i+1,j}(y_{i+1}, \ldots, y_n) := f_{i,j}(l_i(y_{i+1}, \ldots, y_n), y_{i+1}, \ldots, y_n).$$

Step 4. If $i+1 < n$, set $i := i+1$, $\bar{y}^0 := (y^0_{i+1}, \dots, y^0_n)$, and start again with step 2.

Step 5. Consider a first order Taylor development of $F_n = f_{n,n}$ centered at y^0_n, equate it to 0 and call its solution y^1_n.

Step 6. For $i = n-1$ to 1, set $y^1_i := l_i(y^1_{i+1}, \dots, y^1_n)$.

Recall that $y^* \leqslant y^1 \leqslant y^0$ and that $F(y^1) \geqslant 0$ (see [3]). This algorithm can be carried out by assuming that the Jacobian matrix F' is always nonsingular, in which case some kind of pivoting may be necessary. No pivoting has been introduced here, because in the MNT context $F'(y)$ is assumed to be a nonsingular M-matrix and this property is inherited by each F_i, $2 \leqslant i \leqslant n$, as follows by induction from the next two lemmas. As for the former, it is an adaptation of Lemma 3.3 in [5], whose proof omitted a not entirely simple argument included here.

Lemma 3.1. *For each $y \in \langle x^0, y^0 \rangle$, $F'_2(\bar{y})$ is a nonsingular* M-*matrix.*

P r o o f. Notice that

$$l_1(\bar{y}) = y^0_1 - \frac{1}{\partial_1 f_1(y^0)} \left[f_1(y^0) + \sum_{j=2}^{n} \partial_j f_1(y^0)(y_j - y^0_j) \right],$$

so that if $i \neq 1 \neq j$ one gets

$$\begin{aligned}\partial_j f_{2,i}(\bar{y}) &= \partial_j f_i(l_1(\bar{y}), \bar{y}) + \partial_1 f_i(l_1(\bar{y}), \bar{y}) * \partial_j l_1(\bar{y}) \\ &= \partial_j f_i(l_1(\bar{y}), \bar{y}) - \partial_1 f_i(l_1(\bar{y}), \bar{y}) * \frac{\partial_j f_1(y^0)}{\partial_1 f_1(y^0)}.\end{aligned}$$

Since F' is always an M-matrix, hence $F'_2(\bar{y})$ is a Z-matrix, i.e. its off-diagonal terms are nonpositive, because

$$\partial_j f_{2,i}(\bar{y}) \leqslant \partial_j f_i(l_1(\bar{y}), \bar{y}) \leqslant 0 \quad \text{for} \quad j \neq i.$$

With $y \in \langle \bar{x}^0, \bar{y}^0 \rangle$, consider the matrix

$$A := \begin{pmatrix} \partial_1 f_1(y^0) & \dots & \partial_n f_1(y^0) \\ \partial_1 f_2(l_1(\bar{y}), \bar{y}) & \dots & \partial_n f_2(l_1(\bar{y}), \bar{y}) \\ \vdots & \ddots & \vdots \\ \partial_1 f_n(l_1(\bar{y}), \bar{y}) & \dots & \partial_n f_n(l_1(\bar{y}), \bar{y}) \end{pmatrix}.$$

It is apparent that A is a Z-matrix and that $F'(l_1(\bar{y}), \bar{y}) \leqslant A$. Since $F'(l_1(\bar{y}), \bar{y})$ is a nonsingular M-matrix, its associated Jacobi matrix is convergent (see [12]), i.e.

$$r(D(F'(l_1(\bar{y}), \bar{y}))^{-1} * [D(F'(l_1(\bar{y}), \bar{y})) - F'(l_1(\bar{y}), \bar{y})]) < 1,$$

where r denotes the spectral radius and $D(\cdot)$ denotes the corresponding diagonal matrix. By virtue of

$$0 \leqslant D(A)^{-1} * [D(A) - A] \leqslant D(F'(l_1(\overline{y}), \overline{y}))^{-1} * [D(F'(l_1(\overline{y}), \overline{y})) - F'(l_1(\overline{y}), \overline{y})],$$

the Perron-Frobenius theory (see [12]) yields

$$r(D(A)^{-1} * [D(A) - A]) \leqslant r(D(F'(l_1(\overline{y}), \overline{y}))^{-1} * [D(F'(l_1(\overline{y}), \overline{y})) - F'(l_1(\overline{y}), \overline{y})]),$$

whence

$$r(D(A)^{-1} * [D(A) - A]) < 1.$$

Thus A itself is nonsingular and $A^{-1} \geqslant 0$ (see [12]), i.e. it is a nonsingular M-matrix. Consider now

$$M_1 := \begin{pmatrix} 1 & 0 & 0 & \dots & \dots & 0 \\ -m_{2,1} & 1 & 0 & \dots & \dots & 0 \\ -m_{3,1} & 0 & 1 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 1 & 0 \\ -m_{n,1} & 0 & \dots & \dots & 0 & 1 \end{pmatrix} \quad \text{with } m_{i,1} := \frac{\partial_1 f_i(l_1(\overline{y}), \overline{y})}{\partial_1 f_1(y^0)},$$

whence it follows that

$$M_1 * A = \begin{pmatrix} \partial_1 f_1 & \dots & \dots & \partial_n f_1 \\ 0 & & & \\ \vdots & & F_2(\overline{y})' & \\ 0 & & & \end{pmatrix}.$$

Thus $F_2'(\overline{y})$ is nonsingular, and by taking into account the above block structure, one gets

$$A^{-1} * M_1^{-1} = \begin{pmatrix} (\partial_1 f_1)^{-1} & \dots & \dots & \dots \\ 0 & & & \\ \vdots & & (F_2'(\overline{y}))^{-1} & \\ 0 & & & \end{pmatrix}.$$

Since the product on the left-hand side leaves the last $n-1$ columns in A^{-1} unchanged, one gets that $(F_2'(\overline{y}))^{-1} \geqslant 0$, which together with $F_2'(\overline{y})$ being a Z-matrix yields that $F_2'(\overline{y})$ is a nonsingular M-matrix. $\square$

Lemma 3.2. *The following propositions hold:*

(i) F_2' *is isotone on* $\langle \bar{x}^0, \bar{y}^0 \rangle$.

(ii) $0 \leqslant F_2(\bar{y}^0)$.

Proof. (i) It easily follows by adapting Lemma 3.5 in [5].

(ii) See Theorem 3.4 (ii) in [8]. $\square$

Let us now define, for each $1 \leqslant i \leqslant n-1$, and inductively for $j \leqslant i$

$$\begin{aligned} &L_{i,i}(\bar{y}) := l_i(\bar{y}), \quad \text{and} \\ &L_{i,j}(\bar{y}) := l_j(L_{i,j+1}(\bar{y}), L_{i,j+2}(\bar{y}), \ldots, L_{i,i}(\bar{y}), \bar{y}) \quad \text{for } j \neq i. \end{aligned}$$

With this notation it is possible to give a handy representation of $f_{i,i}$ in terms of $f_{1,i} = f_i$, namely

$$f_{i,i}(\bar{y}) = f_i(\widehat{y}) \quad \text{with } \widehat{y} := (L_{i-1,1}(\bar{y}), L_{i-1,2}(\bar{y}), \ldots, L_{i-1,i-1}(\bar{y}), \bar{y}).$$

Notice now that one step of Brown's method can be written in matrix form as

$$T(y^0) * (B(y^0) - y^0) = -\Phi(y^0), \tag{3.1}$$

where $\Phi(y^0) := (f_{i,i}(\bar{y}^0))$. The symbol $T(y^0)$ stands for the upper triangular matrix whose ith row is the gradient $\partial f_{i,i}(\bar{y}^0)$, that is $t_{i,k} := \partial_k f_{i,i}(\bar{y}^0)$ for $i \leqslant k$ and $t_{i,k} := 0$ for $i > k$. Notice also that $T(y^0)$ is the upper triangular part in the LU decomposition of $F'(y^0)$.

In order to analyse the variation of B in K^+, it is necessary to point out that the function Φ itself varies with the initial point y^0; this parametric dependence, which also affects T, will be considered implicit in (3.1), which is all needed here. But when considering the total differential of (3.1), one has to take into account the parametric role of y^0, as well as when considering partial derivatives of the functions involved. This is denoted here by putting the affected functions in round brackets whenever necessary. Thus, from the total differential of (3.1), one gets

$$T(y^0) * \partial B(y^0) = T(y^0) - \partial(\Phi(y^0)) - \partial(T(y^0)) * (B(y^0) - y^0). \tag{3.2}$$

Note that $\partial(T(y^0))$ is a three-linear functional.

Theorem 3.3. *The following inequalities hold componentwise:*

$$\partial(\Phi(y^0)) \leqslant T(y^0) \quad \textit{and} \quad \partial(T(y^0)) \geqslant 0.$$

Proof. The results are a consequence of the following two coupled inductive propositions for $1 \leqslant i \leqslant n-1$:

(a) For $1 \leqslant k \leqslant n$, $1 \leqslant j \leqslant i$, $i+1 \leqslant m$ the functions $L_{i,j}$ satisfy

$$\partial_k(L_{i,j}(\bar{y}^0)) \geqslant \partial_k L_{i,j}(\bar{y}^0) = \partial_k L_{i,j} \geqslant 0, \tag{a.1}$$

and

$$\partial_k(\partial_m L_{i,j}(\bar{y}^0)) \leqslant 0. \tag{a.2}$$

(b) For $1 \leqslant k \leqslant n$ and $j \geqslant i+1$, $f_{i+1,i+1}$ is such that

$$\partial_k(f_{i+1,i+1}(\bar{y}^0)) \leqslant t_{i+1,k}(\bar{y}^0) \tag{b.1}$$

and

$$\partial_k(\partial_j f_{i+1,i+1}(\bar{y}^0)) \geqslant \partial_k \partial_j f_{i+1,i+1}(\bar{y}^0) \geqslant 0. \tag{b.2}$$

Consider first the case $i=1$. As for (a.1), notice that, since

$$l_1(y^0) = y_1^0 - \frac{f_1(y^0)}{\partial_1 f_1(y^0)},$$

it follows that

$$\partial_k(l_1(y^0)) = \delta_{1,k} - \frac{\partial_k f_1(y^0)}{\partial_1 f_1(y^0)} + f_1(y^0) * \frac{\partial^2_{k,1} f_1(y^0)}{(\partial_1 f_1(y^0))^2}, \tag{3.3}$$

where δ stands for the Kronecker symbol. Thus if $k=1$,

$$\partial_1(l_1(\bar{y}^0)) = f_1(y^0) * \frac{\partial^2_{1,1} f_1(y^0)}{(\partial_1 f_1(y^0))^2} \geqslant 0 = \partial_1 l_1,$$

while if $k>1$, we have

$$\partial_k(l_1(\bar{y}^0)) = -\frac{\partial_k f_1(y^0)}{\partial_1 f_1(y^0)} + f_1(y^0) * \frac{\partial^2_{k,1} f_1(y^0)}{(\partial_1 f_1(y^0))^2} \geqslant -\frac{\partial_k f_1(y^0)}{\partial_1 f_1(y^0)} = \partial_k l_1.$$

In order to prove (a.2), recall that

$$\begin{aligned}\partial_k(\partial_m l_1(\overline{y}^0)) &= -\frac{\partial_k(\partial_m f_1(\overline{y}^0))}{\partial_1 f_1(y^0)} + \partial_m f_1(y^0) * \frac{\partial_k(\partial_1 f_1(y^0))}{(\partial_1 f_1(y^0))^2}\\ &= -\frac{\partial^2_{k,m} f_1(\overline{y}^0)}{\partial_1 f_1(y^0)} + \partial_m f_1(y^0) * \frac{\partial^2_{k,1} f_1(y^0)}{(\partial_1 f_1(y^0))^2}.\end{aligned}$$

Since $\partial^2 \geqslant 0$ and $\partial_m f_1 \leqslant 0$ because $m > 1$, (a.2) is finally obtained for $i = 1$.

To prove (b.1), using

$$\begin{aligned}\partial_k(f_{2,2}(\overline{y}^0)) &= \partial_k(f_2(l_1(\overline{y}^0), \overline{y}^0))\\ &= \partial_1 f_2(\hat{y}^0) * \partial_k(l_1(\overline{y}^0)) + (1 - \delta_{1,k}) * \partial_k f_2(\hat{y}^0),\end{aligned}$$

and taking into account (3.3), we conclude that

$$\partial_1(f_{2,2}(\overline{y}^0)) = \partial_1 f_2(\hat{y}^0) * f_1(y^0) * \frac{\partial^2_{1,1} f_1(y^0)}{(\partial_1 f_1(y^0))^2} \leqslant 0.$$

On the other hand, if $k > 1$, then

$$\begin{aligned}\partial_k(f_{2,2}(\overline{y}^0)) &= \partial_1 f_2(\hat{y}^0) * \left(\partial_k l_1 + f_1(y^0) * \frac{\partial^2_{k,1} f_1(y^0)}{(\partial_1 f_1(y^0))^2}\right) + \partial_k f_2(\hat{y}^0)\\ &\leqslant \partial_1 f_2(\hat{y}^0) * \partial_k l_1 + \partial_k f_2(\hat{y}^0) = \partial_k f_{2,2}(\overline{y}^0).\end{aligned}$$

In order to prove now (b.2) with $i = 1$, note first that

$$\begin{aligned}\partial_k(\partial_j f_{2,2}(\overline{y}^0)) &= \partial_k(\partial_1 f_2(\hat{y}^0) * \partial_j l_1 + \partial_j f_2(\hat{y}^0))\\ &= \partial_k(\partial_1 f_2(\hat{y}^0)) * \partial_j l_1 + \partial_k(\partial_j f_2(\hat{y}^0)) + \partial_1 f_2(\hat{y}^0) * \partial_k(\partial_j l_1).\end{aligned}$$

Since from (a.1) one has

$$\begin{aligned}\partial_k(\partial_j f_2(\hat{y}^0)) &= \partial^2_{1,j} f_2(\hat{y}^0) * \partial_k(l_1(\overline{y}^0)) + (1 - \delta_{1,k}) * \partial^2_{k,j} f_2(\hat{y}^0)\\ &\geqslant \partial^2_{1,j} f_2(\hat{y}^0) * \partial_k l_1 + (1 - \delta_{1,k}) * \partial^2_{k,j} f_2(\hat{y}^0),\end{aligned}$$

this inequality and the corresponding one for $j = 1$ yield

$$\begin{aligned}\partial_k(\partial_j f_{2,2}(\overline{y}^0)) \geqslant\ &[\partial^2_{1,1} f_2(\hat{y}^0) * \partial_k l_1 + (1 - \delta_{1,k}) * \partial^2_{k,1} f_2(\hat{y}^0)] * \partial_j l_1\\ &+ [\partial^2_{1,j} f_2(\hat{y}^0) * \partial_k l_1 + (1 - \delta_{1,k}) * \partial^2_{k,j} f_2(\hat{y}^0)]\\ &+ \partial_1 f_2(\hat{y}^0) * \partial_k(\partial_j l_1).\end{aligned} \tag{3.4}$$

On the other hand,

$$\begin{aligned}\partial_k\partial_j f_{2,2}(\overline{y}^0) &= \partial_k(\partial_1 f_2(\hat{y}^0) * \partial_j l_1 + (1-\delta_{1,j}) * \partial_j f_2(\hat{y}^0))\\ &= [\partial^2_{1,1} f_2(\hat{y}^0) * \partial_k l_1 + (1-\delta_{1,k}) * \partial^2_{k,1} f_2(\hat{y}^0)] * \partial_j l_1\\ &\quad + (1-\delta_{1,j})[\partial^2_{1,j} f_2(\hat{y}^0) * \partial_k l_1 + (1-\delta_{1,k}) * \partial^2_{k,j} f_2(\hat{v}^0)].\end{aligned}$$

Since $j > 1$, going back to (3.4) one obtains

$$\partial_k(\partial_j f_{2,2}(\overline{y}^0)) \geqslant \partial_k\partial_j f_{2,2}(\overline{y}^0) + \partial_1 f_2(\hat{v}^0) * \partial_k(\partial_j l_1).$$

Taking into account (a.2) and the inequality $\partial_1 f_2 \leqslant 0$, one finally gets (b.2) for $i = 1$.

Now let us assume that (a) and (b) are valid for all $m < i \leqslant n-1$ with $i > 1$. Hence, for (a.1), it follows that

$$\begin{aligned}\partial_k(L_{i,i}(\overline{y}^0)) = \partial_k(l_i(\overline{y}^0)) &= \partial_k\left(y_i^0 - \frac{f_{i,i}(\overline{y}^0)}{\partial_i f_{i,i}(\overline{y}^0)}\right)\\ &= \delta_{i,k} - \frac{\partial_k(f_{i,i}(\overline{y}^0))}{\partial_i f_{i,i}(\overline{y}^0)} + f_{i,i}(\overline{y}^0) * \frac{\partial_k(\partial_i f_{i,i}(\overline{y}^0))}{(\partial_i f_{i,i}(\overline{y}^0))^2}\\ &\geqslant \delta_{i,k} - \frac{\partial_k(f_{i,i}(\overline{y}^0))}{\partial_i f_{i,i}(\overline{y}^0)}\\ &\geqslant \delta_{i,k} - \frac{\partial_k f_{i,i}(\overline{y}^0)}{\partial_i f_{i,i}(\overline{y}^0)}\\ &= \partial_k l_i = \partial_k L_{i,i}(\overline{y}^0) = \partial_k L_{i,i} \geqslant 0,\end{aligned}$$

where the first inequality above follows from (b.2), while the other from (b.1). This completes the first step in the inner induction from $j = i$ to $j = 1$, in order to prove (a.1). Now, since $L_{i,j}(\overline{y}^0) = l_j(\widehat{y}^0)$ where $\widehat{y}^0 = (L_{i,j+1}(\overline{y}^0), \ldots, L_{i,i}(\overline{y}^0), \overline{y}^0)$, one obtains

$$\begin{aligned}\partial_k(L_{i,j}(\overline{y}^0)) &= \sum_{m=j+1}^{i} \partial_m l_j * \partial_k(L_{i,m}(\overline{y}^0)) + \prod_{m=j+1}^{i}(1-\delta_{m,k}) * \partial_k l_j\\ &\geqslant \sum_{m=j+1}^{i} \partial_m l_j * \partial_k L_{i,m} + \prod_{m=j+1}^{i}(1-\delta_{m,k}) * \partial_k l_j\\ &= \partial_k L_{i,j}(\overline{y}^0) = \partial_k L_{i,j} \geqslant 0\end{aligned}$$

with $\partial_k L_{i,j} = 0$ if $k \leqslant j$. Notice that also the last inequality follows inductively from the inequality for $j = i$.

As for (a.2), another descending induction from $j = i$ to $j = 1$ is performed. Thus for $m \geqslant i+1$, (b.2) yields

$$\begin{aligned} \partial_k(\partial_m L_{i,i}(\bar{y}^0)) &= -\frac{\partial_k(\partial_m f_{i,i}(\bar{y}^0))}{\partial_i f_{i,i}(\bar{y}^0)} + \partial_m f_{i,i}(\bar{y}^0) * \frac{\partial_k(\partial_i f_{i,i}(\bar{y}^0))}{(\partial_i f_{i,i}(\bar{y}^0))^2} \\ &\leqslant -\frac{\partial_k \partial_m f_{i,i}(\bar{y}^0)}{\partial_i f_{i,i}(\bar{y}^0)} \leqslant 0, \end{aligned}$$

while if $j < i$, the (inner) induction and (a.1) as proved for i imply

$$\begin{aligned} \partial_k(\partial_m L_{i,j}(\bar{y}^0)) &= \partial_k\Big(\sum_{s=j+1}^{i} \partial_s l_j * \partial_m L_{i,s}(\bar{y}^0) + \partial_m l_j\Big) \\ &= \sum_{s=j+1}^{i} \partial_k(\partial_s l_j) * \partial_m L_{i,s} \\ &\quad + \sum_{s=j+1}^{i} \partial_s l_j * \partial_k(\partial_m L_{i,s}(\bar{y}^0)) + \partial_k(\partial_m l_j) \leqslant 0. \end{aligned}$$

In order to prove (b.1), notice that from (a.1) it follows as above that

$$\begin{aligned} \partial_k(f_{i+1,i+1}(\bar{y}^0)) &= \sum_{j=1}^{i} \partial_j f_{i+1}(\hat{y}^0) * \partial_k(L_{i,j}(\bar{y}^0)) + \prod_{j=1}^{i}(1-\delta_{k,j}) * \partial_k f_{i+1}(\hat{y}^0) \\ &\leqslant \sum_{j=1}^{i} \partial_j f_{i+1}(\hat{y}^0) * \partial_k L_{i,j} + \prod_{j=1}^{i}(1-\delta_{k,j}) * \partial_k f_{i+1}(\hat{y}^0) \\ &= t_{i+1,k}(\bar{y}^0). \end{aligned}$$

Regarding (b.2), notice first that

$$\begin{aligned} \partial_k(\partial_j f_{i+1,i+1}(\bar{y}^0)) &= \partial_k\Big(\sum_{m=1}^{i} \partial_m f_{i+1}(\hat{y}^0) * \partial_j L_{i,m}(\bar{y}^0) + \partial_j f_{i+1}(\hat{y}^0)\Big) \\ &= \sum_{m=1}^{i} \partial_k(\partial_m f_{i+1}(\hat{y}^0)) * \partial_j L_{i,m} + \partial_k(\partial_j f_{i+1}(\hat{y}^0)) \\ &\quad + \sum_{m=1}^{i} \partial_m f_{i+1}(\hat{y}^0) * \partial_k(\partial_j L_{i,m}(\bar{y}^0)). \end{aligned} \tag{3.5}$$

For each term in (3.5) it is obtained from (a.1) that

$$\begin{aligned}\partial_k(\partial_m f_{i+1}(\hat{y}^0)) &= \sum_{p=1}^{i}\partial^2_{p,m} f_{i+1}(\hat{y}^0) * \partial_k(L_{i,p}(\bar{y}^0)) + \prod_{p=1}^{i}(1-\delta_{k,p}) * \partial^2_{k,m} f_{i+1}(\hat{y}^0)\\ &\geqslant \sum_{p=1}^{i}\partial^2_{p,m} f_{i+1}(\hat{y}^0) * \partial_k L_{i,p} + \prod_{p=1}^{i}(1-\delta_{k,p}) * \partial^2_{k,m} f_{i+1}(\hat{y}^0).\end{aligned}$$

The inequalities in (a) now imply (b.2) as in the case $i = 1$. □

Theorem 3.4. *If $z^0 \in K^+$, then the Brown iterates satisfy $z^k \leqslant y^k$, $k = 1, 2, \ldots$.*

Proof. Theorem 3.3 and (3.2) imply that $\partial B \geqslant 0$ in K^+, because T is always a nonsingular M-matrix. By considering an ascending path g in K^+ joining z^0 to y^0 as in Theorem 2.3, it follows that

$$y^1 - z^1 = B(y^0) - B(z^0) = \int_0^1 \partial_t(B(g(t))\,\mathrm{d}t = \int_0^1 \partial B(g(t)) * \partial g(t)\,\mathrm{d}t \geqslant 0.$$

A simple induction completes the proof. □

4. The analytic Fourier iterations

The Fourier iterations for Brown's method have been introduced in [3] in analogy with the Newton-Fourier iterations. These Brown-Fourier iterates give us a monotone bracketing of y^*, which is moreover contained in the corresponding Newton-Fourier bracketing (see [3]). Their description in the framework of the original Brown algorithm now follows.

Step $1'$. Set $\bar{x}^0 := x^0$ $(i = 1)$ and $F_1^-(x) = (f^-_{1,j}(x)) := (f_j(x)) = F(x)$.

Step $2'$. Consider the affine approximation of $f^-_{i,i}$ centered at $\bar{x}^0$ with the gradient values of $f_{i,i}$ at $\bar{y}^0$ and solve for x_i, i.e. $x_i = l_i^-(x_{i+1}, \ldots, x_n)$.

Step $3'$. Define the $(i+1)$th reduced lower system

$$\begin{gathered}F^-_{i+1}(x_{i+1}, \ldots, x_n) := (f^-_{i+1,j}(x_{i+1}, \ldots, x_n)) = 0 \quad \text{where}\\ f^-_{i+1,j}(x_{i+1}, \ldots, x_n) := f^-_{i,j}(l_i^-(x_{i+1}, \ldots, x_n), x_{i+1}, \ldots, x_n) \quad \text{for } i+1 \leqslant j \leqslant n.\end{gathered}$$

Step $4'$. If $i+1 < n$, set $i := i+1$, $\bar{x}^0 := (x^0_{i+1}, \ldots, x^0_n)$ and start again with step 2'.

Step $5'$. Consider a first order approximation of $f^-_{n,n}$ at x^0_n with the slope given by $\partial_n f_{n,n}(\bar{y}^0)$, equate it to 0 and call its solution x^1_n.

Step 6′. For $i = n-1$ to 1 define $x_i^1 := l_i^-(x_{i+1}^1, \ldots, x_n^1)$.

The x^1 thus obtained is the first Fourier iterate corresponding to the Brown iterate y^1. Recall that $x^0 \leqslant x^1 \leqslant y^*$ and $F(x^1) \leqslant 0$ (see [3]).

Analogously to the previous section, let us now define inductively, for $1 \leqslant i \leqslant n-1$ and for $j \leqslant i$,

$$L_{i,i}^-(\overline{x}) := l_i^-(\overline{x}) = x_i^0 - \frac{1}{\partial_1 f_{i,i}(\overline{y}^0)}\left[f_{i,i}^-(\overline{x}^0) + \sum_{j=i+1}^{n} \partial_j f_{i,i}(\overline{y}^0)(x_j - x_j^0)\right],$$

and for $j \neq i$,

$$L_{i,j}^-(\overline{x}) := l_j^-(L_{i,j+1}^-(\overline{x}), L_{i,j+2}^-(\overline{x}), \ldots, L_{i,i}^-(\overline{x}), \overline{x}).$$

Here one easily gets

$$f_{i,i}^-(\overline{x}) = f_i(\widehat{x}) \quad \text{with } \widehat{x} := (L_{i-1,1}^-(\overline{x}), L_{i-1,2}^-(\overline{x}), \ldots, L_{i-1,i-1}^-(\overline{x}), \overline{x}).$$

Lemma 4.1. *$f_{i,i}^-(\overline{x}^0) \leqslant 0$ for $2 \leqslant i \leqslant n$.*

P r o o f. Consider first $i = 2$. Then

$$\begin{aligned} f_{2,2}^-(\overline{x}^0) &= f_2(\widehat{x}^0) = f_2(L_{1,1}^-(\overline{x}^0), \overline{x}^0) = f_2(l_1^-(\overline{x}^0), \overline{x}^0) - f_2(x^0) + f_2(x^0) \\ &\leqslant \partial_1 f_2(l_1^-(\overline{x}^0), \overline{x}^0) * (l_1^-(\overline{x}^0) - x_1^0) + f_2(x^0) \\ &= \partial_1 f_2(l_1^-(\overline{x}^0), \overline{x}^0) * \left(-\frac{f_1(x^0)}{\partial_1 f_1(y^0)}\right) + f_2(x^0) \\ &\leqslant f_2(x^0) \leqslant 0, \end{aligned}$$

where the first inequality is a consequence of the order convexity. The proof is now completed by induction. Assume that for some i, $2 \leqslant i \leqslant n$,

$$f_{j,j}^-(\overline{x}^0) \leqslant 0 \quad \text{if} \quad 2 \leqslant j \leqslant i-1 < n.$$

Notice first that, because of the order convexity,

$$\begin{aligned} f_{i,i}^-(\overline{x}^0) &= f_i(\widehat{x}^0) = f_i(\widehat{x}^0) - f_i(x^0) + f_i(x^0) \\ &\leqslant \sum_{j=1}^{i-1} \partial_j f_i(\widehat{x}^0) * (L_{i-1,j}^-(\overline{x}^0) - x_j^0) + f_i(x^0). \end{aligned} \tag{4.1}$$

Now it is necessary to prove that

$$L_{i-1,j}^-(\overline{x}^0) - x_j^0 \geqslant 0 \quad \text{for} \quad 1 \leqslant j \leqslant i-1. \tag{4.2}$$

This proof is done by means of an (inner) induction from $i-1$ to 1. Notice first that

$$L^-_{i-1,i-1}(\overline{x}^0) - x^0_{i-1} = -\frac{f^-_{i-1,i-1}(\overline{x}^0)}{\partial_{i-1} f_{i-1,i-1}(\overline{y}^0)} \geqslant 0.$$

For the general term, one has

$$\begin{aligned} L^-_{i-1,j}(\overline{x}^0) - x^0_j &= l_j(\widehat{x}^0) - x^0_j \\ &= -\frac{1}{\partial_j f_{j,j}(\overline{y}^0)} * \left[f^-_{j,j}(\overline{x}^0) + \sum_{k=j+1}^{i-1} \partial_k f_{j,j}(\overline{y}^0)(L^-_{i-1,k}(\overline{x}^0) - x^0_k) \right], \end{aligned}$$

which makes it clear that the inner and outer inductions imply (4.2). By applying these inequalities in (4.1), one finally gets

$$f^-_{i,i}(\overline{x}^0) \leqslant f_i(\overline{x}^0) \leqslant 0,$$

and the proof is complete. □

Denoting $\Phi^-(x^0) := (f^-_{i,i}(\overline{x}^0))$ and $B(x^0, y^0) := x^1$, for the Fourier iterate one has the system

$$T(y^0) * (B(x^0, y^0) - x^0) = -\Phi^-(x^0). \tag{4.3}$$

Let us assume first that x^0 is kept fixed while y^0 is allowed to vary. The corresponding total differential of (4.3) then yields

$$T(y^0) * \partial B(x^0, y^0) = -\partial(\Phi^-(x^0)) - \partial(T(y^0)) * (B(x^0, y^0) - x^0). \tag{4.4}$$

Lemma 4.2. *With the notation as in* (4.4), *it follows that* $\partial(\Phi^-(x^0)) \geqslant 0$.

P r o o f. Note first that $\partial_k(f^-_{1,1}(x^0)) = \partial_k(f_1(x^0)) = 0$, for $1 \leqslant k \leqslant n$.

Since $f^-_{i,i}(\overline{x}^0) \leqslant 0$, $1 \leqslant i \leqslant n$, and by virtue of (b.2) in Theorem 3.3, it is also clear that

$$\begin{aligned} \partial_k(L^-_{i,i}(\overline{x}^0)) = \partial_k(l^-_i(\overline{x}^0)) &= \partial_k\left(x^0_i - \frac{f^-_{i,i}(\overline{x}^0)}{\partial_i f_{i,i}(\overline{y}^0)} \right) \\ &= f^-_{i,i}(\overline{x}^0) * \frac{\partial_k(\partial_i f_{i,i}(\overline{y}^0))}{(\partial_i f_{i,i}(\overline{y}^0))^2} \leqslant 0. \end{aligned}$$

It now inductively follows, from $j = i$ to $j = 1$, that

$$\begin{aligned} \partial_k(L^-_{i,j}(\overline{x}^0)) &= \sum_{m=j+1}^{i} \partial_m l^-_j * \partial_k(L^-_{i,m}(\overline{x}^0)) \\ &= -\sum_{m=j+1}^{i} \frac{\partial_m f_{j,j}(\overline{y}^0)}{\partial_j f_{j,j}(\overline{y}^0)} * \partial_k(L^-_{i,m}(\overline{x}^0)) \leqslant 0. \end{aligned}$$

This inequality yields

$$\partial_k(f^-_{i+1,i+1}(\overline{x}^0)) = \partial_k(f_{i+1}(\widehat{x}^0)) = \sum_{j=1}^{i} \partial_j f_{i+1}(\widehat{x}^0) * \partial_k(L^-_{i,j}(\overline{x}^0)) \geqslant 0,$$

because $\partial_j f_{i+1} \leqslant 0$ for $j \leqslant i$, which completes the proof. $\square$

Theorem 4.3. *If $z^0 \in K^+$, then $B(x^0, y^0) \leqslant B(x^0, z^0)$.*

Proof. Lemma 4.2 and Theorem 3.3 applied in (4.4) imply that $\partial B(x^0, y) \leqslant 0$ for y in K^+. The conclusion then follows as in Theorem 3.4. $\square$

Suppose now that y^0 is held fixed while x^0 varies. In this case, from the corresponding differential in (4.3), it is easy to see that

$$T(y^0) * \partial B(x^0, y^0) = T(y^0) - \partial(\Phi^-(x^0)). \tag{4.5}$$

Lemma 4.4. *With the notation as in* (4.5) *it follows that* $\partial(\Phi^-(x^0)) \leqslant T(y^0)$.

Proof. Clearly one has that

$$\partial_k(f_1(x^0)) = \partial_k f_1(x^0) \leqslant \partial_k f_1(y^0) = t_{1,k}(y^0).$$

Assuming now for some $i < n-1$ and all k, $1 \leqslant k \leqslant n$, that

$$\partial_k(f^-_{i,i}(\overline{x}^0)) \leqslant t_{i,k}(y^0) = \partial_k f_{i,i}(y^0) = t_{i,k}(y^0),$$

we obtain that

$$\begin{aligned}
\partial_k(l^-_i(\overline{x}^0)) = \partial_k(L^-_{i,i}(\overline{x}^0)) &= \partial_k\left(x^0_i - \frac{f^-_{i,i}(\overline{x}^0)}{\partial_i f_{i,i}(\overline{y}^0)}\right) \\
&= \delta_{i,k} - \frac{\partial_k(f^-_{i,i}(\overline{x}^0))}{\partial_i f_{i,i}(\overline{y}^0)} \\
&\geqslant \partial_k l^-_i = \partial_k l_i \geqslant 0,
\end{aligned}$$

so that

$$\begin{aligned}
\partial_k(L^-_{i,j}(\overline{x}^0)) &= \partial_k(l^-_j(\widehat{x}^0)) \\
&= \sum_{m=j+1}^{i} \partial_m l^-_j * \partial_k(L^-_{i,m}(\overline{x}^0)) + \prod_{m=j+1}^{i} (1-\delta_{m,k}) * \partial_k(l^-_j) \\
&\geqslant \sum_{m=j+1}^{i} \partial_m l^-_j * \partial_k L^-_{i,m} + \prod_{m=j+1}^{i} (1-\delta_{m,k}) * \partial_k l^-_j \\
&= \sum_{m=j+1}^{i} \partial_m l_j * \partial_k L_{i,m} + \prod_{m=j+1}^{i} (1-\delta_{m,k}) * \partial_k l_j \\
&= \partial_k L^-_{i,j}(\overline{y}^0) = \partial_k L_{i,j} \geqslant 0.
\end{aligned}$$

Since one has

$$\partial_m f_{i+1}(\widehat{x}^0) \leqslant \partial_m f_{i+1}(\widehat{y}^0) \leqslant 0,$$

for $1 \leqslant m \leqslant i$, one obtains

$$\begin{aligned}\partial_k(f^-_{i+1,i+1}(\overline{x}^0)) &= \partial_k(f_{i+1}(\widehat{x}^0))\\ &= \sum_{m=1}^{i} \partial_m f_{i+1}(\widehat{x}^0) * \partial_k(L^-_{i,m}(\overline{x}^0)) + \prod_{m=1}^{i}(1-\delta_{k,m}) * \partial_k f_{i+1}(\widehat{x}^0)\\ &\leqslant \sum_{m=1}^{i} \partial_m f_{i+1}(\widehat{y}^0) * \partial_k L_{i,m} + \prod_{m=1}^{i}(1-\delta_{k,m}) * \partial_k f_{i+1}(\widehat{y}^0)\\ &= t_{i+1,k}(y^0),\end{aligned}$$

which completes the proof. □

Theorem 4.5. *If $w^0 \in K^-$, then $B(x^0, y^0) \leqslant B(w^0, y^0)$.*

Proof. Lemma 4.4, when applied in (4.5), implies that $\partial(B(x, y^0)) \geqslant 0$ for x in K^-. Since K^- is an order convex set, the ascending segment joining x^0 to w^0 is contained in K^-, so that the argument in the proofs of Theorems 3.3 and 4.3 applies again and yields the conclusion. □

Corollary 4.6. *If $w^0 \in K^-$ and $z^0 \in K^+$, then $B(x^0, y^0) \leqslant B(w^0, z^0)$.*

5. An example

An illustration is briefly discussed in this section. It deals with Chandrasekhar's equation and suggests that, given the quadratic convergence of Brown's analytic method, one should only expect a modest improvement in the number of iterations by conveniently choosing the starting (upper) point.

Consider thus Chandrasekhar's equation, namely

$$v(t) = 1 - \frac{1}{4}\int_0^1 \Big(\frac{t}{s+t} * \frac{1}{v(s)}\Big)\,\mathrm{d}s, \quad 0 \leqslant t \leqslant 1.$$

The approach to dealing with this equation follows [4] and [8]. For $h := \frac{1}{64}$, the trapezoidal integration rule is applied at the points ih, $0 \leqslant i \leqslant 64$. Taking into account that $v(0) = 1$, the resulting nonlinear system is

$$F(x) := (f_i(x)) = 0, \quad 1 \leqslant i \leqslant 64,$$

where

$$f_i(x) := x_i + \frac{1}{4}\left[w_0 + \sum_{j=1}^{64} w_j * \frac{i}{i+j} * \frac{1}{x_j}\right] - 1,$$

$$w_0 := w_{64} := \frac{h}{2} \quad w_j := h, \quad 1 \leqslant j \leqslant 63.$$

Consider two different starting upper points y^0 whose coordinates are all equal to 1 in one case and to 5 in the other. It is easily verified that $F(y^0) \geqslant 0$ in both cases. Consider also the corresponding Fourier iterations with $x_i^0 := .5$, $1 \leqslant i \leqslant 64$, for which it is also easy to see that $F(x^0) \leqslant 0$. The stopping criteria for Brown and Fourier iterations are given by the first k for which the function residues satisfy, respectively,

$$\|F(y^k)\|_\infty < \varepsilon := .5 * 10^{-13}, \quad \text{and} \quad \|F(x^k)\|_\infty < \varepsilon.$$

The computations have been carried out with the double precision of Fortran 77. The table shows the values of the iterates approximating $v(1)$, namely y_{64}^k, as well as the values of x_{64}^k. The exact digits are underlined. The final k in each column of values is the one for which the function values satisfy the corresponding stopping criterion given above.

k	y_{64}^k	x_{64}^k	y_{64}^k	x_{64}^k
0	5.	.5	1.	.5
1	.808462758084	$.\underline{7}897145052 00$	$.\underline{799}636685607$	$.\underline{79}3434227609$
2	$.\underline{799}218390107$	$.\underline{7991}26316604$	$.\underline{799194}762887$	$.\underline{7991}84364766$
3	$.\underline{7991947027}34$	$.\underline{79919470}0358$	$.\underline{799194702574}$	$.\underline{7991947025}44$
4	$.\underline{799194702574}$	$.\underline{799194702574}$		$.\underline{799194702574}$

It is worth pointing out not only the consistency of the table with the results in the paper, but also that Brown iterates appear to converge faster than the corresponding Fourier iterates. This is a well established fact by Ostrowski when $n = 1$, namely for one dimensional analytic Newton-Fourier iterations (see [10]).

Final comment. *The extension of the results in the paper to discretized Brown iterations will be analyzed elsewhere, along with their possible application to the comparison of discretized Brown and Newton iterations.*

Acknowledgement. The referee's constructive criticism has been very helpful for the improvement of the manuscript.

References

[1] *R. Brent*: Some efficient algorithms for solving systems of nonlinear equations. SIAM J. Numer. Anal. *10* (1973), 327–344.

[2] *K. Brown*: A quadratically convergent Newton-like method based upon Gaussian elimination. SIAM J. Numer. Anal. *6* (1969), 560–569.

[3] *A. Frommer*: Monotonicity of Brown's method. Z. Angew. Math. Mech. *68* (1988), 101–109.

[4] *A. Frommer*: Comparison of Brown's and Newton's method in the monotone case. Numer. Math. *52* (1988), 511–521.

[5] *J. P. Milaszewicz, S. Abdel Masih*: Elimination and fixed point iterations. Comput. Math. Appl. *25* (1993), 43–53.

[6] *J. P. Milaszewicz*: Comparison theorems for monotone Newton-Fourier iterations and applications in functional elimination. Linear Algebra Appl. *220* (1995), 343–357.

[7] *J. P. Milaszewicz*: Comparison theorems for a third order method. Appl. Math. Lett. *10* (1997), 17–21.

[8] *J. P. Milaszewicz*: On Brown's and Newton's methods with convexity hypotheses. J. Comput. Appl. Math. *150* (2002), 1–24.

[9] *J. M. Ortega, W. C. Rheinboldt*: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York-London, 1970.

[10] *A. M. Ostrowski*: Solution of Equations and Systems of Equations. Academic Press, New York-London, 1960.

[11] *H. Schwetlick*: Numerische Lösung Nichtlinearer Gleichungssysteme. R. Oldenburg Verlag, München, 1979.

[12] *R. S. Varga*: Matrix Iterative Analysis. Prentice-Hall, Englewood Cliffs, 1962.

Author's address: *J. P. Milaszewicz*, Instituto Argentino de Matemática, CONICET, Saavedra 15, 1083 Buenos Aires, Argentina, e-mail: `milas@mate.dm.uba.ar`.